Multi-Frequency Trade*

Nicolas Crouzet, Ian Dew-Becker, and Charles G. Nathanson

March 23, 2017

Abstract

We develop a noisy rational expectations model of financial trade featuring investors who acquire information and trade at a range of different frequencies. In the model, a restriction on high-frequency trading lowers the efficiency of prices at high frequencies but increases the efficiency of prices at low frequencies. In a particular equilibrium of the model, investors specialize into trading at individual frequencies. High- and low-frequency investors coexist, trade with each other, and make money from each other. The model matches numerous basic features of financial markets: investors endogenously specialize into strategies distinguished by frequency; volume is disproportionately driven by high-frequency traders; and the portfolio holdings of informed investors forecast returns at the same frequencies as those at which they trade.

Investors in financial markets follow many different strategies, including value investing, technical analysis, macro strategies, and algorithmic trading. These strategies differ in two salient ways. First, they require investors to learn about different aspects of asset prices; market-makers or algorithmic traders care more about the high-frequency movements of prices, while value investing puts more emphasis on their slow-moving features. These investors all understand that their information sets may not overlap, and yet they trade with each other, presumably making some money in the process. Second, these strategies differ in the frequency at which they require investors to trade, or equivalently the rate at which they turn over their positions.

This paper proposes an equilibrium model in which investors endogenously specialize in acquiring information and trading at different frequencies. There is a single fundamentals process, and a continuum of investors who trade forward contracts on the fundamental. These investors also learn about different aspects of asset dynamics. An example of the fundamentals process is the spot price of oil: investors are able to acquire information that tells them about the future path of oil prices, allowing them to potentially earn profits on the forward contracts. As is common elsewhere, in order to grease the wheels of the market, we assume that investors trade against an exogenous flow of demand for forward contracts that fluctuates stochastically over time.

*Crouzet: Northwestern University. Dew-Becker: Northwestern University and NBER. Nathanson: Northwestern University. We appreciate helpful comments from Stijn Van Nieuwerburgh, Ioanid Rosu, and seminar participants at Northwestern and the Adam Smith conference.

We show that in such a model, there exists a natural (though not necessarily unique) equilibrium in which individual investors endogenously choose to focus on specific frequencies of the fundamentals. Some investors learn about low-frequency aspects of oil prices in the sense that they get a signal about their average path over, say, a period of decades, while others learn about higher-frequency behavior, receiving a signal about how oil prices vary from day to day or month to month. This occurs despite the fact that the learning technology is fully general, and in no way tilts investors towards frequency specialization ex-ante.

Given attention allocation – what aspect of fundamentals investors choose to learn about – in equilibrium we show that their positions fluctuate at the frequency at which they receive signals. That is, investors who learn about long-run fundamentals hold positions in forward contracts that fluctuate slowly over time, whereas those who do high-frequency research have positions that vary at high frequencies. Investors choose to learn about high- or low-frequency aspects of fundamentals, and that learning causes them to endogenously become high- or low-frequency traders.

While there is other research on investors who trade at different frequencies, that work typically endows investors with investment horizons that differ exogenously.[1] In our setting, all investors have the same objective, maximizing utility over identical horizons. We view this as an important restriction in our setting because it is obviously not the case that people who trade at high frequencies, e.g. turning over their portfolios once per day, really have investment horizons of only 24 hours. Rather, all investors want to maximize the same utility function over wealth, they just go about it in different ways.

What is particularly interesting about the equilibrium that we obtain is that it is not the case that the informed investors trade only with the exogenous demand (i.e. liquidity traders). In fact, high- and low-frequency traders trade with each other. The simple reason is that a high-frequency trader cannot distinguish uninformative demand shocks from the orders of informed low-frequency traders (and vice versa). So in periods when fundamentals are persistently strong, low-frequency traders tend to hold persistently more forward contracts than high-frequency traders and earn profits from them. Similarly, if there is a very transitory increase in fundamentals, the high-frequency traders tend take advantage and earn profits while the low-frequency traders lose, as they ignore the temporary trading opportunity. In that sense, then, everybody is a noise trader sometimes, and they all understand that, but they still participate and make money on average.

The model has a number of predictions for observable features of financial markets. First, as we have already discussed, it predicts that there are traders who can be distinguished by the frequencies at which their asset holdings change over time, and they do research about fundamentals at the same frequencies. So we obtain endogenous high- and low-frequency traders with a specific prediction for how research aligns with trade. Second, the model also matches salient facts about

[1]See, e.g., Amihud and Mendelson (1986), who assume that investors are forced to sell after random periods of time; Hopenhayn and Werner (1996), who assume that investors vary in their rates of pure time preference, and Defusco, Nathanson, and Zwick (2016) who assume that there are sets of investors who are exogenously forced to sell at determinstic horizons that vary across groups. Turley (2012), like us, studies a setting in which investors endogenously choose to learn about high- or low-frequency information, though he studies only a two-period case.

differences in volume across investors. We can very easily show analytically that high-frequency investors account for a fraction of aggregate volume that is out of proportion to their fraction of total asset holdings.

The model is fundamentally about differences in information across investors. People obtain information in order to make money, and so their asset holdings in general should forecast returns. We see that both in the model and in the data.[2] But different investors' holdings do not forecast returns in the same way. The frequency at which an investor's portfolio holdings forecasts returns is the same as the frequency at which they trade: high-frequency investors' positions forecast returns at very short horizons, while buy-and-hold investors' portfolios forecast returns over much longer periods.

The idea that an investors' asset holdings should forecast returns over a period related to how long those assets will be held is perhaps not surprising. Studies of the holdings of mutual funds and other institutional investors typically examine returns over a period of perhaps 3–12 months. At the other extreme, Brogaard, Hendershott, and Riordan (2014) show that the holdings of high-frequency traders (as defined by NASDAQ) forecast returns over periods of 1–5 seconds – a horizon 7 orders of magnitude smaller than a calendar quarter.

To empirically test our model, we provide novel evidence on the relationship between turnover and asset return predictability. Using form 13F data on institutional asset holdings, we first show that asset turnover within funds is highly persistent over time, suggesting that it is a salient feature of investor strategies. Next, after confirming past results that institutional holdings predict returns, we show that the predictive power of the holdings of high-turnover funds decays much more quickly than those of low-frequency funds, consistent with the model.

Finally, we use the model to study the effects of a policy that restricts high-frequency trade. This policy lowers the informativeness of prices about high-frequency dividend information and increases the high-frequency volatility in returns when it is binding (in the sense that high-frequency traders exist without the policy). However, the policy can *increase* the informativeness of prices about low-frequency dividend information like the mean of dividends over time. This result hinges on a reallocation of investor attention away from high-frequency trading to low-frequency trading; this reallocation occurs when investors devote a fixed stock of attention to trading, but not when the opportunity cost of attention remains constant. By allowing investor specialization to be endogenous, our model sheds light on welfare-relevant trade-offs that are relevant for regulating high-frequency trading.

In general models with dynamic trade are extremely difficult to solve; solutions typically require some kind of restriction, such as to a very narrow class of driving processes (e.g. AR(1) and Ornstein–Uhlenbeck processes studied in Wang (1993, 1994) and He and Wang (1995)), or to only two or three period horizons. We allow for a very long investment horizon and place only technical constraints on the fundamental processes driving the model and obtain fully analytic solutions. The

[2] See, e.g., the literature on the predictive power of mutual fund and institutional investor asset holdings for future returns, such as Carhart (1997) and Yan and Zhang (2009), among many others.

sacrifice that we make is that information sets are fixed on date 0 – investors only obtain signals once. The model should be thought of as essentially a stationary equilibrium: it gives a steady-state description of trade, volume, and returns. It is not well suited to studying how investors and markets respond to shocks to information sets.

The major advantage of our particular information structure is that it allows us to take a long-horizon dynamic model and solve it as a series of parallel scalar problems. In particular, solving our model is only marginally more difficult than solving a standard single-period/single asset noisy rational expectations model – it reduces to a parallel set of such equilibria. The paper thus has useful methodological contributions for analyzing models of trade over time.

This paper builds on a growing recent literature that tries to understand optimal information acquisition in financial markets. The most important building blocks are the models of Van Nieuwerburgh and Veldkamp (2010) and Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016) in that we use a highly similar information and market structure and build on their results on optimal information acquisition (their reverse water-filling solution, in particular). Those papers themselves build critically on work by Grossman and Stiglitz (1980), Hellwig (1980), Diamond and Verrecchia (1981), and Admati (1985) on rational expectations equilibria. More recently, research has tried to understand the effects on the equilibria developed in those earlier papers of various limits on information gathering ability (e.g. Banerjee and Green (2015) and Dávila and Parlatore (2016)).

There is also a literature on price dynamics in rational expectations equilibria, though it is relatively small given how difficult dynamic models are to solve. In particular, a series of papers by Wang (1993, 1994) and He and Wang (1995) study the implications of dynamic equilibria on prices and volume. Those papers are based around AR(1) or Ornstein–Uhlenbeck-type dynamics to maintain tractability (see also Wachter (2002)), whereas we study a setting in which the various exogenous time series may follow processes with minimally constrained autocorrelations. Furthermore, we focus on how investment strategies differ across investors, whereas those papers focus on symmetric strategies. A number of papers also study overlapping generations models, which can help alleviate some of the difficulties with dynamic trade.[3]

There is also a large literature on disagreement in financial markets. In addition to the above work, see, e.g., Townsend (1983), Scheinkman and Xiong (2003), Basak (2005), Hong and Stein (2007), and Banerjee and Kremer (2010), who focus, like us, on dynamics. In our setting, disagreement arises not just because agents receive signals that have random errors, but also because their signals have different relationships with fundamentals. High-frequency and low-frequency investors will often disagree about the price path of the asset over time because they learn about different characteristics of fundamentals – the path over the next few minutes, say, versus the path over the next several years.

Our desire to develop a model that can match salient features of the cross-section of investment strategies follows from a large empirical literature that documents the behavior of many different types of investors and how it affects the aggregate behavior of financial markets. Chen, Jegadeesh,

[3]See Spiegel (1998), Watanabe (2008), and Banerjee (2011), among others.

and Wermers (2000), Gompers and Metrick (2001), Nagel (2005), Griffin and Xu (2009), Yan and Zhang (2009), and Brogaard, Hendershott, and Riordan (2014), for example, study the behavior of institutional investors and how their holdings relate to asset returns. Turley (2012), Bai, Philippon, and Savov (2016), and Weller (2016) study how price informativeness has changed over time and how it is affected by trading costs and the number of investors who trade at different frequencies.

Finally, our work is related to a small literature that studies the properties of asset returns and portfolio choice in the frequency domain including Bandi and Tamoni (2014), Chinco and Ye (2016), Chaudhuri and Lo (2016), and Dew-Becker and Giglio (2016).

The remainder of the paper is organized as follows. Section 1 describes the basic environment, and we solve for optimal information acquisition in section 2. Section 3 examines the implications of the model for the behavior of individual investors in a setting that features investors who specialize in trade at a particular frequency. Section 4 presents empirical evidence on the behavior of institutions consistent with out model of specialization. Finally, section 5 presents our key results on the effects of restrictions on high-frequency trade on return volatility and price efficiency at different frequencies, and section 6 concludes.

1 Asset market equilibrium

We begin by describing the basic market structure and the asset market equilibrium. This section introduces the description of trading strategies in terms of frequencies and shows how the frequency transformation makes multi-period investment a purely scalar problem. the problem is solved from the perspective of date 0.

1.1 Market structure

Time is denoted by $t \in \{-1, 0, 1, ..., T\}$, with T even, and we will focus on cases in which T may be treated as large. There is a fundamentals process D_t that investors make bets on with realizations on all dates except -1 and 0. The time series process is stacked into a vector $D \equiv [D_1, D_2, ..., D_T]'$ (variables without subscripts denote vectors) and is distributed as

$$D \sim N(0, \Sigma_D). \tag{1}$$

The fundamentals process is assumed to be stationary, meaning that it has constant unconditional autocovariances. Stationarity implies that Σ_D is Toeplitz (all diagonals are constant), and we further assume that the eigenvalues of Σ_D are finite and bounded away from zero.[4]

On date 0, there is a market for forward claims on D_t for all $t > 0$. A unit mass of investors indexed by $i \in [0, 1]$ meets on date 0 and commits to a set of trades of futures contracts maturing on all dates. P_t denotes the price of a claim to the fundamental D_t.

[4] The analysis is similar if a transformation of D_t (e.g. its first difference) is stationary. See appendix section A.

There is an exogenous supply of futures, Z, which is distributed as

$$Z \sim N(0, \Sigma_Z). \tag{2}$$

Z_t may be thought of as either exogenous liquidity demand or noise trading. The time series process for supply is also assumed to be stationary. For markets to clear, the net demand of the investors for the fundamental on date t must equal Z_t,[5]

$$\forall t: \int_i Q_{i,t} di = Z_t, \tag{3}$$

where $Q_{i,t}$ is the number of date-t forward claims agent i buys.

A concrete example of a potential process D_t is the price of crude oil: oil prices follow some stochastic process and investors trade futures on oil at many maturities. D_t can also be interpreted as the dividend on a stock, in which case P_t is the price of a forward claim on a single dividend.

1.1.1 Modeling equities

While the concept of a futures market on the fundamentals will be a useful analytic tool, we can also obviously price portfolios of futures. We model equity as a claim to the stream of fundamentals over time. To purchase such a claim, one would enter into futures contracts for the fundamental on each date $t+j$. Since futures contracts specify that money only changes hands at maturity, the money that must be set aside on date t for a futures contract that expires at $t+j$ is $P_{t+j}R^{-j}$, where R is the discount rate (which is assumed here to be a constant). The date-t cost of a claim to the entire future stream of fundamentals is therefore

$$P_t^{equity} \equiv \sum_{j=1}^{T-t} R^{-j} P_{t+j} \tag{4}$$

Holding any given combination of futures claims on the fundamental D is therefore also equivalent to holding futures contracts on equity claims (i.e. committing to a trading strategy in equities at prices that are agreed on at date 0). Any desired set of exposures to fundamentals over time can be obtained either through purchases of futures or through suitable trading strategies for the equity claim (assuming prices can be committed to or that they are predetermined, which will be the case in our equilibrium).

Our analysis of pricing will focus on futures as they will give the most direct analog to past work. When we discuss volume and trading costs, though, we will take advantage of the equity-based implementation.

[5] It is also possible to assume that there is an exogenous downward-sloping supply curve of the fundamental that shifts stochastically over time; our results go through similarly. This case is treated as part of the analysis of appendix 6.

1.2 Information structure

The realization of the time series of fundamentals, $\{D_t\}_{t=1}^T$, can be thought of as a single draw from a multivariate normal distribution. Investors are able to acquire signals about that realization. The signals are a collection $\{Y_{i,t}\}_{t=1}^T$ *observed on date 0* with

$$Y_{i,t} = D_t + \varepsilon_{i,t},\ \varepsilon_i \sim N(0, \Sigma_i), \tag{5}$$

Information sets in the model are fixed on date 0. Through $Y_{i,t}$, investors can learn about fundamentals potentially arbitrarily far into the future. $\varepsilon_{i,t}$ is a stationary error process in the sense that $cov(\varepsilon_{i,t}, \varepsilon_{i,t+j})$ depends on j but not t (again, Σ_i is Toeplitz).

The information structure here is obviously stylized. One interpretation is that we are collapsing to date 0 all the realizations of a stationary process. That is, agents have a machine that gives them signals about fundamentals plus an error, and that machine reports all of its output on date 0. The information structure is meant to generate two important features in the model. First, obviously on any particular date agents can choose to learn about fundamentals on more than just a single date in the future – they can potentially get information about fundamentals in many different periods (e.g. next quarter vs. over the next five years). Second, by restricting $\varepsilon_{i,t}$ to be "stationary", we are forcing agents to choose a fixed policy for information. They build a machine (or a research department) that, rather than yielding information about only a single date, returns information about the entire fundamentals stream over time in a way that places no particular emphasis on any single date.

In choosing Σ_i agents will have two choices to make. First, they will be able to choose how informative their signals are by choosing the variance of ε_i. Second, though, they will be able to choose how accurate the signals are about fundamentals over different horizons. some choices for Σ_i will yield signals that are informative about transitory variation in fundamentals, while others will yield signals that are more useful for forecasting trends.

1.3 Investment objective

All trading decisions are made on date 0. Investors choose demands $\{Q_{i,t}\}_{t=1}^T$ conditional on their observed signals, $\{Y_{i,t}\}_{t=1}^T$, and the set of futures prices, $\{P_t\}_{t=1}^T$. That is, as in past work, agents submit to a central auctioneer demand curves that condition on prices.

We assume that investors have mean-variance utility over cumulative excess returns. Investor i's objective is

$$U_{0,i} = \max_{\{Q_{i,t}\}} E_{0,i}\left[T^{-1}\sum_{t=1}^{T} Q_{i,t}(D_t - P_t)\right] - (\rho T)^{-1} Var_{0,i}\left[\sum_{t=1}^{T} Q_{i,t}(D_t - P_t)\right], \tag{6}$$

where $E_{0,i}$ is the expectation operator conditional on agent i's date-0 information set, $\{P, Y_i\}$. $Var_{0,i}$ is the variance operator conditional on $\{P, Y_i\}$. ρ is risk-bearing capacity per unit of time.

The assumption that all plans are made on date 0 only restricts information sets in a very specific way: it means that investors are not able to condition demand on the *realized* history of fundamentals. That is, it is not free to condition on the history of D_t, even when that history has already been realized. Instead, what agents must condition on is noisy signals about fundamentals, $Y_{i,t}$.

An important implication of that assumption is that agents have no desire to change their investment choices after date 0 since they receive no further information. Agents' trading strategies can thus be equivalently implemented either through a set of purchases of futures contracts or through a dynamic plan for trading the equity claim, as described above.

We interpret the objective as representing a target that an institutional investor might have. Rather than aiming to maximize the discounted sum of returns, as a person who consumes out of wealth might, the investors we study maximize a measure of their performance. The objective can be thought of as representing CARA or quadratic preferences over the sum of excess returns, so it would appear if a manager were paid on date T a fee proportional to total excess returns up to that time. Bhattacharya and Pfleiderer (1985) and Stoughton (1993) also argue that a quadratic contract (which would induce mean-variance preferences) can appear optimally in delegated investment problems. The important characteristic of (6) is that it yields a stationary problem in the sense that there is no discounting to make returns in some periods more important than others.

Finally, note that all investors have the same investment horizon. We show in appendix F that the investment horizon as defined here by T has no effect on information choices in the model – two investors with different T will be equally likely to be high- or low-frequency investors. The simplest way to confirm that fact is to simply note, when we obtain the equilibrium strategies, that T has no effect on the type of information that investors optimally obtain.

1.4 Equilibrium

Conditional on the information choices of the agents – that is, taking the set of Σ_i (which may differ across agents) as given – we study a standard asset market equilibrium.

Definition 1 *An asset market equilibrium is a set of demand functions, $Q_i(P, Y_i)$, and a vector of prices, P, such that investors maximize utility, $U_{0,i}$, and all markets clear, $\int_i Q_{i,t} di = Z_t \; \forall t$.*

The equilibrium concept is that Grossman and Stiglitz (1980), Hellwig (1980), Diamond and Verrecchia (1981), and Admati (1985). Investors submit demand curves for each futures contract to a Walrasian auctioneer who selects equilibrium prices to clear all markets.

The structure is in fact mathematically that of Admati (1985), who studies investment across a set of assets that might represent stocks in different companies, and the solution from that paper applies directly here. Here we are considering investment across a set of futures contracts that represent claims on some fundamentals process across different dates. We simply rotate the Admati (1985) structure from a cross-section to a time series.

1.5 Trading frequencies

This paper is fundamentally about the behavior of markets at different frequencies, so we need a rigorous concept of what frequencies are. We use the fact that fluctuations at different frequencies represent an (asymptotic) orthogonal decomposition of any time series.

Define a set of $T \times 1$ vectors of cosines and sines at the fundamental frequencies $\omega_j = 2\pi j/T$ for $j \in \left\{0, 1, ..., \frac{T}{2}\right\}$

$$c_j \equiv \sqrt{\frac{2}{T}} \left(\cos\left(2\pi j \frac{(t-1)}{T} \right) \right)_{t=1}^{T} \tag{7}$$

$$s_j \equiv \sqrt{\frac{2}{T}} \left(\sin\left(2\pi j \frac{(t-1)}{T} \right) \right)_{t=1}^{T} \tag{8}$$

A cycle at frequency ω_j has an associated wavelength $2\pi/\omega_j$. $\omega_0 = 0$ thus corresponds to an infinite wavelength, or a permanent shock (a constant vector). ω_1 corresponds to a cycle that lasts as long as the sample – c_1 is a single cycle of a cosine. $\omega_{\frac{T}{2}} = \pi$, the highest frequency, corresponds to a cycle that lasts two periods, so that $c_{\frac{T}{2}}$ oscillates between $\pm\sqrt{2/T}$.

The frequency-domain counterpart to the vector of fundamentals, D, is then

$$d = \Lambda' D \tag{9}$$

$$\text{where } \Lambda \equiv \left[\frac{1}{\sqrt{2}} c_0, c_1, ..., c_{\frac{T}{2}-1}, \frac{1}{\sqrt{2}} c_{\frac{T}{2}}, s_1, s_2, ..., s_{\frac{T}{2}-1} \right]. \tag{10}$$

We use the notation $d_j = c_j' D$ and $d_{j'} = s_j' D$ to refer to fundamentals at particular frequencies. When the distinction is necessary, we use the notation j to refer to a frequency associated with a cosine transform and j' to refer to one with a sine transform. In what follows, lower-case letters denote frequency-domain objects. Note that Λ is orthonormal with $\Lambda' = \Lambda^{-1}$.

To understand the above formulas, consider the simple example in which $T = 2$. In this case, the low-frequency component of dividends is $d_0 = (D_1 + D_2)/\sqrt{2}$ and the high-frequency component of dividends is $d_1 = (D_1 - D_2)/\sqrt{2}$. Investors "trade" the low-frequency component d_0 by buying an equal amount of the claims on D_1 and D_2. Converse, investors "trade" the high-frequency component d_1 by buying offsetting amounts of the claims on D_1 and D_2.[6] In general, because d is a linear function of D, it can be thought of as a vector of payoffs on portfolios of futures given by Λ – portfolios with weights on D_t that fluctuate over time as sines and cosines.

For our purposes, the key feature of Λ is that it approximately diagonalizes *all* Toeplitz matrices and thus orthogonalizes stationary time series.[7]

[6] One explicit example of this type of trade is the calendar spread future, through which investors buy the difference in the payoff of a future on different dates. See Cuny (2006) for more information on these contracts.

[7] This is a textbook result that appears in many forms, e.g. Shumway and Stoffer (2011). Brillinger (1981) and Shao and Wu (2007) give similar statements under weaker conditions.

Definition 2 *f_X is the* spectrum *of X with elements $f_{X,j}$, defined as*

$$f_{X,j} \equiv \sigma_{X,0} + 2\sum_{s=1}^{T-1} \sigma_{X,j} \cos(\omega_j s) \quad (11)$$

$$f_X \equiv \left[f_{X,0}, f_{X,1}, ..., f_{X,\frac{T}{2}-1}, f_{X,\frac{T}{2}}, f_{X,1}, f_{X,2}, ..., f_{X,\frac{T}{2}-1} \right]'. \quad (12)$$

Lemma 1 *For a stationary time series $X_t \sim N(0, \Sigma_X)$ with autocovariances $\sigma_{X,j} \equiv cov(X_t, X_{t-j})$,*

$$x \equiv \Lambda' X \Rightarrow N(0, diag(f_X)) \quad (13)$$

where $\Rightarrow$ denotes convergence in the sense that

$$\left| \Lambda' \Sigma_X \Lambda - diag(f_X) \right| \leq c_X T^{-1/2} \quad (14)$$

for a constant c_X and for all T.[8] *$diag(f_X)$ is a matrix with the vector f_X on its main diagonal and zero elsewhere.*

Proof. This is a textbook result (e.g. Brockwell and Davis (1991)). See appendix B for a derivation specific to our case. ■

For any finite horizon, the matrix Λ does not exactly diagonalize the covariance matrix of D. But as T grows, the error induced by ignoring the off-diagonal elements of the covariance matrix of $\Lambda' D$ becomes negligible (it is of order $T^{-1/2}$), and x is well approximated as a vector of independent random variables.[9] The spectrum of X, f_X, measures the variance in X coming from fluctuations at each frequency. It also represents an approximation to the eigenvalues of Σ_X.[10]

To see why this lemma is useful, consider the vector of fundamentals in the frequency domain, $d = \Lambda' D$. Given that $D \sim N(0, \Sigma_D)$, where Σ_D is Toeplitz, we have

$$\Lambda' D = d \Rightarrow N(0, diag(f_D)). \quad (15)$$

Λ thus approximately diagonalizes the matrix Σ_D, meaning that the elements of d – the fluctuations in fundamentals at different frequencies (with both sines and cosines) are jointly asymptotically independent. Moreover, the same matrix Λ asymptotically diagonalizes the covariance matrix of *any* stationary process. That result will allow us to massively simplify the study of investment over

[8] Two technical points may be noted here. First, as a technical matter, the spectrum f_X must be extended as T grows. A simple way to do that is to suppose that there is a true process for X with a spectrum that is a continuous function f_X, and in any finite sample of length T, there is then an associated spectrum $f_{X,T}$ defined in (11). The second point is that the constant c_X is then a function of that true spectrum f_X; the appendix elaborates on that fact.

[9] For all the stationary processes studied in the paper, we assume that the autocovariances are summable in the sense that $\sum_{r=1}^{\infty} |j\sigma_{X,j}|$ is finite (which holds for finite-order stationary ARMA processes, for example).

[10] f_X represents an approximation to the eigenvalues only in the sense that $\Lambda' \Sigma_X \Lambda \approx diag(f_X)$. Providing a sense in which f_X is actually close to the true eigenvalues of Σ_X is a subtler problem that we do not address here. The specific result in lemma 1 is all that we actually need for our results.

many horizons. It says that set of orthogonal factors underlying all stationary processes is (nearly) the same.

1.6 Market equilibrium in the frequency domain

The approximate diagonalization induced by Λ allows us to solve the model through a series of parallel scalar problems that can be easily solved by hand. Using the asymptotic approximation that d and z are independent across frequencies (and across sines and cosines), we obtain the following frequency-by-frequency solution to the asset market equilibrium.[11]

Solution 1 *Under the approximations $d \sim N(0, diag(f_D))$ and $z \sim N(0, diag(f_Z))$, the prices of the frequency-specific portfolios, p_j, satisfy, for all j, j'*

$$p_j = a_{1,j} d_j - a_{2,j} z_j \quad (16)$$

$$a_{1,j} \equiv 1 - \frac{f_{D,j}^{-1}}{\left(\rho f_{avg,j}^{-1}\right)^2 f_{Z,j}^{-1} + f_{avg,j}^{-1} + f_{D,j}^{-1}} \quad (17)$$

$$a_{2,j} \equiv \frac{a_{1,j}}{\rho f_{avg,j}^{-1}} \quad (18)$$

$$\text{where } f_{avg,j}^{-1} \equiv \int_i f_{i,j}^{-1} di \quad (19)$$

where p_j, d_j, and z_j represent the frequency-j components of prices, fundamentals, and supply, respectively. f_i is the spectrum of the matrix Σ_i. See appendix C for the derivation.

The price of the frequency-j portfolio depends only on fundamentals and supply at that frequency. As usual, the informativeness of prices, through $a_{1,j}$, is increasing in the precision of the signals that investors obtain, while the impact of supply on prices is decreasing in signal precision and risk tolerance. The frequency domain analog to the usual demand function is

$$q_{i,j} = \rho \frac{E\left[d_j - p_j \mid y_{i,j}, p_j\right]}{Var\left[d_j - p_j \mid y_{i,j}, p_j\right]}. \quad (20)$$

These solutions for the prices and demands are the standard results for scalar markets. What is novel here is that the choice problem refers to trades over time. p_j is the price of a portfolio whose exposure to fundamentals fluctuates over time at frequency $2\pi j/T$. Both prices and demands at frequency j depend only on signals and supply at frequency j – the problem is completely separable across frequencies.

The appendix shows that the frequency domain solution provides a close approximation to the true solution in the time domain. Specifically, the true time domain solution from Admati (1985)

[11] A simple way to see where this solution comes from is to note that, under the asymptotic approximation, $\Lambda' A_1 \Lambda$ and $\Lambda' A_2 \Lambda$ from the Admati solution can be written purely in terms of diagonal matrices, for which addition, multiplication, and inversion are simply scalar operations on the main diagonal.

(with no approximations) can be written as

$$P = A_1 D - A_2 Z \tag{21}$$

for a pair of matrices A_1 and A_2 defined in the appendix that are complicated matrix functions of Σ_Z, Σ_D, and the precisions of the signals agents obtain.

Proposition 1 *The difference between calculating the prices directly in the time domain using the Admati (1985) solution in the time domain and rotating the frequency domain solution back into the time domain is small in the sense that*

$$\left|A_1 - \Lambda diag\left(a_1\right)\Lambda'\right| \leq c_1 T^{-1/2} \tag{22}$$
$$\left|A_2 - \Lambda diag\left(a_2\right)\Lambda'\right| \leq c_2 T^{-1/2} \tag{23}$$

for constants c_1 and c_2. Furthermore, while prices and demands are stochastic, the time- and frequency-domain solutions are related through an even stronger result

$$e_{\max}\left[Var\left(\Lambda p - P\right)\right] \leq c_P T^{-1/2} \tag{24}$$
$$e_{\max}\left[Var\left(\Lambda q_i - Q_i\right)\right] \leq c_Q T^{-1/2} \tag{25}$$

where the operator $e_{\max}[\cdot]$ denotes the maximum eigenvalue of a matrix (that is, the operator norm), for constants c_P and c_Q.

In other words, among portfolios whose squared weights sum to 1, the maximum variance of the pricing and demand errors – the difference between the truth from the time domain solution and the frequency-domain approximation that assumes that Λ diagonalizes the covariance matrices – is of order $T^{-1/2}$ (that is, the bound holds for *any* portfolio of futures, not just the frequency- or time-domain claims). We note also that these are not limiting results – they are true for all T.

Result 1 shows that for large T, the standard time-domain solution for stationary time series processes becomes arbitrarily close to a simple set of parallel scalar problems in the frequency domain. The time domain solution is obtained from the frequency domain solution by premultiplying by Λ.

2 Optimal information choice

We now model a constraint on information acquisition and characterize optimal strategies. The objective, constraint, and solution are drawn from Van Nieuwerburgh and Veldkamp (2009) and Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016; KVNV). Our analysis follows theirs closely, except that we are studying a time-series model and a frequency transformation. Whereas KVNV study a symmetric equilibrium in which all investors follow the same information acquisition strategy, we will subsequently argue for the relevance of a separating equilibrium in our setting.

2.1 Objective

Following KVNV, we assume that investors choose information to maximize the expectation of their mean-variance objective (6) subject to a linear constraint on total precision:

$$\max_{\{f_{i,j}\}} E_{-1}\left[U_{i,0} \mid \Sigma_i^{-1}\right] \text{ such that } tr\left(\Sigma_i^{-1}\right) \leq \bar{f}^{-1} \tag{26}$$

where E_{-1} is the expectation operator on date -1, i.e. prior to the realization of signals and prices (as distinguished from $E_{i,0}$, which conditions on P and Y_i). The trace function $tr\left(\Sigma_i^{-1}\right)$ measures the total cost of acquiring a private signal with precision matrix Σ_i^{-1} and is subject to the bound $\bar{f}^{-1}$.[12] This cost function is also equal to the sum of the eigenvalues of the precision matrix. Since the eigenvalues represent the precisions of the orthogonalized signals, it can be thought of as measuring the total precision of the independent parts of the signals. Moreover, since the trace operator is invariant under rotations, this measure of information is invariant to the domain of analysis, time or frequency.[13] That is,

$$tr\left(\Sigma_i^{-1}\right) = \sum_{j,j'} f_{i,j}^{-1} \tag{27}$$

The information constraint is linear in the frequency-specific precisions. Investors also face the constraint that $f_{i,j} = f_{i,j'}$, which ensures that the variance matrix of ε_i is symmetric and Toeplitz.[14]

The constraint in (26) implies that investors possess a fixed amount of attention, $\bar{f}^{-1}$, that they may allocate only to learning about the dividend stream. In reality, professional investors may allocate their scarce attention to activities outside of finance as well. To capture this extensive margin of reallocating attention outside of finance, we use the alternative formulation of the objective function given by

$$\max_{\{f_{i,j}\}} E_{-1}\left[U_{i,0} \mid \Sigma_i^{-1}\right] - \theta tr\left(\Sigma_i^{-1}\right), \tag{28}$$

where $\theta > 0$ captures the opportunity cost of attention. The distinction between (26) and (28) matters only in Section 5, where we explore comparative statics. We use (26) going forward unless noted otherwise.

The appendix shows that, given the optimal demands, an agent's expected utility is linear in the precision they obtain at each frequency.

Lemma 2 *Under the frequency domain representation, when informed investors optimize, each*

[12]Our main analysis considers the case where signals about fundamentals are costly but investors can condition on prices freely. Appendix J considers a case where it is costly to condition expectations on prices and shows that model's predictions results go through similarly with the caveat that investors never choose to become informed about prices, as in Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016).

[13]This result relies on the approximation $\Sigma_i \approx \Lambda' diag\left(f_i\right)\Lambda$.

[14]KVNV show that the solution of the optimal attention allocation problem (26) are identical if one assumes that the cost of information is measured by the entropy of the investor's signals, which corresponds to the function $\ln\left|\Sigma_i^{-1}\right| \approx \sum_{j,j'} \log f_{i,j}^{-1}$. The key feature of the two cost functions is that they are non-convex in precision.

investor's expected utility may be written as a function of their own precisions, $f_{i,j}^{-1}$, and the average across other investors, $f_{avg,j}^{-1} \equiv \int_i f_{i,j}^{-1} di$, with

$$E_{-1}\left[U_{0,i} \mid \{f_{i,j}\}\right] = \frac{1}{2T} \sum_{j,j'} \lambda_j \left(f_{avg,j}^{-1}\right) f_{i,j}^{-1} + constants \tag{29}$$

$\lambda_j(x)$ is a function determining the marginal benefit of information at each frequency with the properties $\lambda_j(x) > 0$ and $\lambda_j'(x) < 0$ for all $x \geq 0$. The fact that $\lambda_j' < 0$ says that the marginal benefit to an investor of allocating attention to frequency j is decreasing in the amount of attention that other investors allocate to that frequency – attention decisions are strategic substitutes. If $f_{avg,j}^{-1}$, the average precision of the signals obtained by other agents, is high, then prices are already efficient at frequency j, so there is little benefit to an investor from learning about that frequency.

The frequency-domain transformation is what allows us to write utility as a simple sum across frequencies. An investor's utility depends additively on the amount of information that they obtain at each frequency. In the time domain, utility is a complicated function of matrices.

2.2 Characterizing the optimum

The critical feature of (29) is that expected utility is linear in the set of precisions that agent i chooses, $\left\{f_{i,j}^{-1}\right\}$. Since the both the objective (29) and the constraint (27) are linear in the choice variables, it immediately follows that agents either allocate all attention to a single frequency, or that they are indifferent between allocating attention across some subset of the frequencies. We then obtain the following solution for attention allocation.

Solution 2 *Information is allocated so that*

$$f_{avg,j}^{-1} = \begin{cases} \lambda_j^{-1}\left(\bar{\lambda}\right) & \text{if } \lambda_j(0) \geq \bar{\lambda} \\ 0 & \text{otherwise} \end{cases} \tag{30}$$

where $\bar{\lambda}$ is obtained as the solution to

$$\sum_{j,j':\lambda_j(0)>\bar{\lambda}} \lambda_j^{-1}\left(\bar{\lambda}\right) = \bar{f}^{-1}. \tag{31}$$

This is the reverse water-filling solution from KVNV. While it may appear mathematically complicated, the intuition is simple: investors allocate attention to signals in such a way that the marginal benefit is equalized to the extent possible across frequencies. It is impossible to allocate negative attention, though, so if the marginal benefit of paying zero attention to a particular frequency, $\lambda_j(0)$, is below the cutoff $\bar{\lambda}$, then $f_{i,j}^{-1} = 0$ there for all investors.

The intuition is easiest to develop graphically. Figure 1 plots the functions $\lambda_j(0)$ and $\lambda_j\left(f_{avg,j}^{-1}\right)$ across frequencies ω_j, where

$$\lambda_j(0) = f_{D,j}\left(1 + \rho^{-2} f_{D,j} f_{Z,j}\right). \tag{32}$$

The initial marginal benefit of allocating attention is increasing in the amount of fundamental information and the volatility of supply.

The details of the calibration are reported in appendix K. What is important here is simply that $\lambda_j(0)$ has peaks at low, middle, and high frequencies. Those are the frequencies at which D_t or Z_t is more volatile, so there is more information to potentially be gathered and a larger reward for doing so. For a given value of $\bar{f}^{-1}$, $\lambda_j\left(f_{avg,j}^{-1}\right)$ is a flat line for all j such that $\lambda_j(0) \geq \bar{\lambda}$. Those are the frequencies that investors learn about. The term "reverse water filling" refers here to the idea that the curve $\lambda_j(0)$ is inverted and one pours water into it. $\bar{\lambda}$ is then the level of the water's surface.[15] As the information constraint is relaxed, $\bar{\lambda}$ falls and potentially more frequencies receive attention.

Given the calibration, we see that there are investors acquiring information in three disconnected ranges of frequencies. At the places where $\lambda_j(0)$ is farther above $\bar{\lambda}$, there is more information acquisition, whereas the locations where $\lambda_j(0) = \bar{\lambda}$ are marginal in the sense that they are the next to receive attention if $\bar{\lambda}$ falls.

Another way to interpret the results is to observe the following:

Result 1 *The return at frequency j has variance*

$$Var[r_j] = \lambda_j\left(f_{avg,j}^{-1}\right) \quad (33)$$

$$\text{where } r_j \equiv d_j - p_j. \quad (34)$$

The marginal benefit of acquiring information at a particular frequency is exactly equal to the unconditional variance of returns at that frequency. When returns have high variance, there are potentially large profits to be earned from acquiring information. When returns have zero variance, on the other hand, prices are already perfectly informative, so there is no reason to study fundamentals at such a frequency. So agents desire to learn at the frequencies where returns are most volatile.

The solution derived here characterizes aggregate information acquisition – the sum of the precisions obtained by all the agents at each frequency – but it does not describe exactly what strategy each agent follows; and in fact there are infinitely many strategies for individual investors consistent with the aggregate solution. We now examine one particular solution that leads to the existence of traders who can be characterized by their trading frequencies.

[15] Again, each frequency (except 0 and π) has an associated sine and cosine. The same amount of precision is required to be allocated to both the sine and cosine at each frequency.

3 Specialization

3.1 The model with specialization

Given the assumptions we have made so far, the only restrictions on information allocation are those that ensure that the information allocation condition (30) holds. There are numerous equilibria with that characteristic, though. KVNV focus on the symmetric equilibrium in which all investors allocate their attention in proportion to $f_{avg,j}^{-1}$ at each frequency. There are also asymmetric and mixed-strategy equilibria. To put it differently, in any equilibrium, the *aggregate* allocation of attention across frequency is always given by (30); however, *individual's* attention allocation is not pinned down.

Since one of our goals is to understand the potential existence and behavior of high-and low-frequency traders, we now focus on equilibria in which all investors learn about only a single frequency. Specifically, we assume that for every agent i, there is a frequency j_i^* such that:

$$f_{i,j} = f_{i,j'} = \begin{cases} \bar{f}^{-1}/2 \text{ if } j = j_i^* \text{ and } j_i^* \notin \left\{0, \frac{T}{2}\right\} \\ \bar{f}^{-1} \text{ if } j = j_i^* \text{ and } j_i^* \in \left\{0, \frac{T}{2}\right\} \\ 0 \text{ otherwise} \end{cases} \tag{35}$$

($\bar{f}^{-1}$ is divided by 2 in the first case because the agent pays attention to both the sine and the cosine at frequency j_i^*). Specialization here means that agents obtain information about a single frequency and are uninformed about all other frequencies.

Why would focusing on this class of equilibria be natural? A simple reason could be that learning about a new frequency involves fixed costs. The following lemma, the proof of which is reported in appendix G, establishes that, in that case, the unique equilibria are those in which agents specialize in learning about only one frequency.

Lemma 3 *Assume that the attention allocation constraint is:*

$$tr(f_i^{-1}) + \kappa \sum_{j,j'} \mathbb{1}_{\left\{f_{i,j}^{-1} > 0\right\}} \leq \bar{f}^{-1}, \tag{36}$$

where $0 < \kappa < \bar{f}^{-1}$ is a fixed cost of learning about each frequency. Let $\bar{\lambda}$ be the unique solution to:

$$\sum_{j,j' \ s.t. \ \lambda_j(0) < \bar{\lambda}} \lambda_j^{-1}(\bar{\lambda}) = \bar{f}^{-1} - \kappa.$$

Then the unique equilibria are those where for each j such that $\bar{\lambda} > \lambda_j(0)$, a fraction γ_j of agents only learn about frequency j, where:

$$\gamma_j = \frac{\lambda_j^{-1}(\bar{\lambda})}{\bar{f}^{-1} - \kappa}.$$

The intuition for the lemma is straightforward. Since in equilibrium, all frequencies pay the same marginal benefit, learning about more than one frequency requires an extra payment of the fixed cost with zero benefit. As a result, individuals specialize in learning about only one frequency.[16] The condition $\bar{f}^{-1} \geq \bar{f}_{min}^{-1} + \kappa$ guarantees that The lemma additionally shows that this is true even if the fixed cost is vanishingly small.

A natural question about this result is: how should one interpret the fixed cost of learning about a frequency in the time domain? The following lemma adresses this issue formally.

Lemma 4 *The information allocation constraint on frequencies is equivalent to the information allocation constraint on time-series signals:*

$$tr(\Sigma_i^{-1}) + \kappa \lim_{\alpha \to +\infty} g_\alpha(\Sigma_i^{-1}) \leq \bar{f}^{-1},$$

(when that limit exists), where the function g_α *is given by* $g_\alpha(M) = tr\left[\left(I - e^{-2\alpha M}\right)\left(I + e^{-2\alpha M}\right)^{-1}\right]$. *In particular, if* Σ_i *is diagonal (that is, in the case of independent time-series signals), then the constraint on frequencies is equivalent to:*

$$\sum_{t=0}^{T-1} \tau_{i,t} + \kappa \sum_{t=0}^{T-1} \mathbb{1}_{\{\tau_{i,t}>0\}} \leq \bar{f}^{-1},$$

where $\tau_{i,t}$ *denotes the* $t-th$ *diagonal entry of* Σ_i^{-1}*, i.e. the precision of the time-series signal about date-t dividends.*

Since $tr(\Sigma_i^{-1}) = tr(f_i^{-1})$, this lemma says that the analog of fixed costs of learning about specific frequencies, in the time domain, is the term $\lim_{\alpha \to +\infty} g_\alpha(\Sigma_i^{-1})$. As discussed in appendix, this terms represents a penalization of the *number* of non-zero eigenvalues of Σ_i^{-1}. Thus, one can think of the constraint (36) as requiring agents to pay a fixed cost for every independent component they add to their the signal structure $\{\epsilon_{i,t}\}$. In the limit where signals are time-independent, this analogy is exact, and the constraint (36) implies that agents pay a fixed cost for each time-t signal with non-zero precision.

Now obviously in reality nobody learns about just a single aspect of the world. It is also not the case, though, that everybody learns about everything. We focus here on the case with specialization as it is consistent with the evidence discussed above and with new results presented below on the wide divergences in behavior and research across investors.

3.2 Specialization model predictions

We now examine the implications of the model with specialization for the behavior of individual investors, obtaining the following results:

[16]Note that the equilibria are not unique: agents remain indifferent between learning about one frequency or another. However, they strictly prefer allocating their attention to one frequency, than splitting it equally across frequencies.

1. Investors can be distinguished by the frequencies at which their portfolio positions fluctuate, and those fluctuations match the frequencies at which they obtain information.

2. The average volume accounted for by an investor is proportional to the frequency at which they trade. In the presence of quadratic trading costs, costs can be linearly decomposed across frequencies and are quadratic in frequency.

3. Investors' positions are correlated with returns most strongly at the frequency they learn about.

4. Investors earn money from liquidity provision, they earn money from trading at the frequency at which they are informed, and they *lose* money to other investors from trading at frequencies at which they are uninformed.

3.2.1 Fluctuations in positions

Result 2 *Investor i's demand at frequency j is*

$$q_{i,j} = z_j + \rho \left[\left(f_{i,j}^{-1} - f_{avg,j}^{-1} \right) r_j + f_{i,j}^{-1} \tilde{\varepsilon}_{i,j} \right] \tag{37}$$

where $\tilde{\varepsilon}_{i,j}$ is equal to the jth column of Λ multiplied by ε_i, i.e. the noise in investor i's signal at frequency j, and r_j is the realized return on the jth frequency portfolio.

Investor i's demand depends on three terms. z_j is the stochastic supply at frequency j. Each investor is equally willing to absorb supply, so they all take equal fractions, giving them a common component z_j.

The second term, $\rho \left(f_{i,j}^{-1} - f_{avg,j}^{-1} \right) r_j$ reflects investor i's information. At the frequency that investor i pays attention to, $f_{i,j}^{-1} - f_{avg,j}^{-1}$ is positive, so investor i's demand covaries positively with returns at that frequency. That is, investors who learn about low-frequency dynamics hold portfolios that are long when returns are high over long periods, while high-frequency investors hold portfolios that covary positively with transitory fluctuations in returns. At the other frequencies, where investor i does not pay attention, $f_{i,j}^{-1} = 0$, so the investor's demand actually covaries slightly negatively with returns, holding z_j fixed.

The third term, $\rho f_{i,j}^{-1} \tilde{\varepsilon}_{i,j}$ is the idiosyncratic part of demand that is due to the random error in the signal that agent i receives. Note that the standard deviation of $f_{i,j}^{-1} \tilde{\varepsilon}_{i,j}$ is equal to $f_{i,j}^{-1/2}$, so these errors are equal to zero at the frequencies that the investor ignores (i.e. all but one).

When the number of active frequencies (i.e. with $f_{avg,j}^{-1} > 0$) is large, $f_{avg,j}^{-1}$ becomes small relative to $\bar{f}^{-1}$. That means that the term $\left(f_{i,j}^{-1} - f_{avg,j}^{-1} \right)$ is close to zero at all frequencies except for the one that the agent pays attention to, j_i^*. Since $\left(f_{i,j}^{-1} - f_{avg,j}^{-1} \right) \approx 0$ for all other frequencies, we have

$$Q_{i,t} \approx Z_t + \cos\left(\omega_{j_i^*} t/T \right) \left(\bar{f}^{-1} r_{j_i^*} + \tilde{\varepsilon}_j \right) + \sin\left(\omega_{j_i^*} t/T \right) \left(\bar{f}^{-1} r_{j_i^{*\prime}} + \tilde{\varepsilon}_{j'} \right) \tag{38}$$

Investor i's demand on date t thus is approximately equal to supply on that date plus a multiple the part of returns depending on frequency $\omega_{j_i^*}$, $r_{j_i^*}$ and $r_{j_i^{*\prime}}$, plus an error. The second line shows that what is really going on is that investor i's information can be thought of as a signal about returns interacted with a cosine and a sine.

The important feature of equations (37) and (38) is that they show that each agent's position is equal to Z_t plus fluctuations that come primarily at the frequency that they pay attention to.[17] That is, if some agent allocates all attention to frequency $\omega_{j_i^*}$, then their relative position, $Q_{i,t} - Z_t$, fluctuates over time at frequency $\omega_{j_i^*}$. This can be seen by noting that the sum of a sine and a cosine at frequency $\omega_{j_i^*}$, even with different coefficients, remains a cosine that fluctuates at frequency $\omega_{j_i^*}$, just shifted by a constant. Specifically,

$$Q_{i,t} \approx Z_t + \sqrt{\begin{array}{c} \left(\bar{f}^{-1} r_{j_i^*} + \tilde{\varepsilon}_{j_i^*}\right)^2 \\ + \left(\bar{f}^{-1} r_{j_i^{*\prime}} + \tilde{\varepsilon}_{j_i^{*\prime}}\right)^2 \end{array}} \cos\left(\omega_{j_i^*} t/T + C_{i,j}\right) \tag{40}$$

where $C_{i,j}$ is a function of $\left(\bar{f}^{-1} r_j + \tilde{\varepsilon}_j\right)$ and $\left(\bar{f}^{-1} r_{j'} + \tilde{\varepsilon}_{j'}\right)$. So agent i's excess demand is approximately a cosine with a random translation and amplitude.

As a numerical example, figure 2 plots a hypothetical history for a particular agent's position $Q_{i,t}$ in the same calibration that we studied above. We see that $Q_{i,t}$ looks like a sinusoid with noise added; the noise is from the Z_t term in (40). The noise in the agent's signal, $\tilde{\varepsilon}_{j_i^*}$ and $\tilde{\varepsilon}_{j_i^{*\prime}}$, simply changes the amplitude and translation of the cosine in (40).

So equations (37) and (38) deliver our first two basic results for the behavior of individual specialized investors: the investors can be distinguished by the frequencies at which their asset holdings fluctuate, and those frequencies are linked to the type of information that they acquire. The first result, that there are traders at different frequencies, is essentially obtained by design: it follows from the assumption that agents specialize across frequencies. Nevertheless, the finding is interesting for its novelty in a theoretical setting.

The fact that the frequency of trading is related to information acquisition, while not surprising, is certainly not obtained by assumption. In past work, different trading behavior has sometimes been obtained by simply assuming that different agents have different exogenously specified trading horizons. In our case, any investor can potentially trade at any frequency. That choice is entirely endogenous – investors are not forced to trade any particular frequency by assumption (the assumption is that they gather information at a single frequency).

The reason that buy-and-hold investors in our model buy and hold is that they have persistent low-frequency information about fundamentals – they have signals that fundamentals will be strong

[17] More formally, the variance of $Q_{i,t} - Z_t$ can be decomposed as $Var\left(Q_{i,t} - Z_t\right) = \sum_j \left(\rho^2 \left(f_{i,j}^{-1} - f_{avg,j}^{-1}\right)^2 f_{R,j} + f_{i,j}^{-1}\right)$. Now consider a simple case where there are N frequencies that receive equal allocations of information. Furthermore, denote the spectrum of returns as $f_{R,j}$. Then we have

$$\lim_{N \to \infty} Var\left(Q_{i,t} - Z_t\right) = \rho^2 \bar{f}^{-2} f_{R,j_i^*} + f_{i,j}^{-1} \tag{39}$$

which shows that $Q_{i,t} - Z_t$ is driven by fluctuations at a single frequency.

or weak over long time spans. Similarly, high-frequency investors have transitory high-frequency information. So the model provides a testable prediction that we should observe investors doing research about asset return dynamics that aligns in terms of frequency or time horizon with their average holding periods.

3.2.2 How do investors earn money?

Investors earn returns in the model through two basic mechanisms: providing liquidity and trading on private signals. We can see from the results on demand above that the liquidity function is spread equally across investors. The effects of private information are more interesting.

Result 3 *Investor i's expected profits (which are also equal to the covariance of their positions with returns) are*

$$E\left[Q_i'R\right] = \sum_j E\left[q_{i,j}r_j\right] \quad (41)$$

$$= \sum_j E\left[z_j r_j\right] + \rho\left(\bar{f}^{-1} - f_{avg,j_i^*}^{-1}\right) f_{R,j_i^*} - \sum_{j \neq j_i^*} \rho f_{avg,j}^{-1} f_{R,j} \quad (42)$$

where the spectrum of returns is (from (33))

$$f_{R,j} = \max\left(\bar{\lambda}, \lambda_j(0)\right) \quad (43)$$

The first term on the right-hand side is the contribution from each investor's liquidity provision. The second term is the positive covariance of the investor's holdings with returns at the frequency they are informed about. In informed investors have demands that covary positively with returns at a particular frequency, then the investors who are uninformed about that frequency must have demands that covary *negatively* with returns (after accounting for $E\left[z_j r_j\right]$). That is the third term above: there is a negative contribution to the correlation of the investor's demands with returns from the frequencies they do not pay attention to.

It is not the case that trading from frequencies $j \neq j_i^*$ is unprofitable. Investors still earn profits from liquidity provision. It is just the case that some of their profits at those frequencies are taken by investors who are more informed. In some sense, this result is inevitable. The total profits that the informed investors earn as a group come from trading with liquidity demand. If an investor earns more money by becoming informed at some frequency, that must come at the cost of other investors.

Now since the information allocation we find is an optimum, obviously investors must be in some sense comfortable with the losses we see here. Intuitively, the slight trading losses they bear at frequencies other than j_i^* are offset by their gains at j_i^*. But obviously any trading that informed investors do that is not related to exogenous supply must ultimately come at the cost of other informed investors.

So the model has the feature that high-frequency investors earn money at high frequencies, but they lose money at lower frequencies relative to other investors. Low-frequency investors might know that oil prices are on a long-term downward trend. In such a situation, the high-frequency investors can still earn profits by betting on day-to-day movements in oil prices, but they will lose money to those who understand that prices are generally drifting down. Similarly, low-frequency investors will tend to lose out at high frequencies by, for example, failing to trade at precisely the right time, buying slightly too high and selling slightly too low compared to where they would if they had high-frequency information.

3.2.3 Volume and trading costs

We study volume in the representation of the model in terms of equity holdings. Recall that equity is modeled as a discounted claim to dividends on all future dates. An investor's position $Q_{i,t}$ can be acquired either by holding $Q_{i,t}$ units of forwards or $Q_{i,t}$ units of equity. In modeling volume, we consider trading in equity. Using equity to measure volume ensures that a person who has position that does not change between dates t and $t+1$ ($Q_{i,t} = Q_{i,t+1}$) induces no trade volume, whereas if we assumed that every forward position required volume, then each investor's contribution would be $|Q_{i,t}|$ each date, meaning, unrealistically, that buy-and-hold investors would contribute constantly to volume.

The equity volume contributed by investor i is

$$V_{i,t} = |\Delta Q_{i,t}| \tag{44}$$

$$\text{where } \Delta Q_{i,t} \equiv Q_{i,t} - Q_{i,t-1} \tag{45}$$

Recall that investors' positions can be written as functions of cosines and sines. The appendix derives the following result for volume for each investor.

Result 4 *The volume induced by investor i, $|\Delta Q_{i,t}|$, may be approximated as*

$$|\Delta Q_{i,t}| \approx |\Delta Z_t| + \omega_{j_i^*} \bar{f}^{-1} \rho \left| \begin{array}{c} \sin\left(\omega_{j_i^*} t\right)\left(r_{j_i^*} + \tilde{\varepsilon}_{i,j_i^*}\right) \\ +\cos\left(\omega_{j_i^*} t\right)\left(r_{j_i^{*\prime}} + \tilde{\varepsilon}_{i,j_i^{*\prime}}\right) \end{array} \right| \tag{46}$$

and has expectation

$$E\left[|\Delta Q_{i,t}|\right] - E\left[|\Delta Z_t|\right] \approx \omega_{j_i^*} \sqrt{\frac{2}{\pi}} \rho \left(\bar{f}^{-1}\bar{\lambda} + 1/2\right). \tag{47}$$

The approximations converge to true equalities as $T \to \infty$.

So we find that agent i's contribution to volume depends on the volume induced by exogenous supply and also the magnitude of returns at frequency $\omega_{j_i^*}$ ($\bar{\lambda}$).

Agent i's contribution to aggregate volume is also exactly proportional to the frequency they allocate attention to, $\omega_{j_i^*}$. High-frequency investors contribute relatively more to aggregate volume because they have portfolios than change most rapidly. An investor at the very lowest frequency, $\omega = 0$, contributes zero volume beyond that induced by exogenous supply, since their position is nearly constant. Investors at $\omega = \pi$, on the other hand, contribute the maximum possible volume as they approximately turn over their entire portfolios in each period.

Not surprisingly, it is also straightforward to show that high-frequency investors typically will face larger trading costs. As an example, consider quadratic trading costs proportional to $\sum_{t=2}^{T} (Q_{i,t} - Q_{i,t-1})^2$. The appendix shows that trading costs can, just like volume, be decomposed across frequencies.

Result 5 *The quadratic variation in an investor's position can be approximated (with convergence as $T \to \infty$) by*

$$\sum_{t=2}^{T} (Q_{i,t} - Q_{i,t-1})^2 = \sum_{j,j'} (2\pi j)^2 T^{-1} q_{i,j}^2. \tag{48}$$

The quadratic trading costs associated with a given demand vector Q_i can be written as a simple sum across frequencies. Trading costs are proportional to the frequency squared. It is thus immediately apparent from our frequency-domain analysis that changes in the magnitude of trading costs have the largest effects on the highest frequencies.

4 Institutional portfolio turnover and return forecasting

In this section, we demonstrate empirically that investment funds specialize in the frequency at which they trade, and we show that the portfolio holdings of high turnover funds forecast returns at relatively shorter horizons than those of lower turnover funds.

4.1 Data

We obtain data on institutional asset holdings from SEC form 13F. These forms list the identities and quantities of securities held by each institution at the end of the filing quarter.[18] The data cover the period 1980–2015. Data on monthly stock returns is taken from CRSP and is aggregated to a quarterly frequency with delisting returns included. We obtain data on the risk-free rate, market return, and Fama–French (1993) factors from Kenneth French's website.

[18] Institutions are required to report only their holdings of 13(f) securities, a category defined by the SEC that includes exchange-traded equities and some securities that can be converted to equity. Only institutions holding more than $100,000,000 in 13(f) securities at the end of the quarter must file form 13F, and each institution is required to report only securities for which its holdings exceed $200,000 or 10,000 shares. Gompers and Metrick (2001) provide more information on these filings. We use Thompson Reuters's database of these filings, which includes the price of each security at the filing date.

4.2 Fund specialization

Yan and Zhang (2009) define the churn rate $c_{i,t}$ of institution i in quarter t as

$$c_{i,t} \equiv \frac{\min\left(\sum_{s|\Delta S_{i,s,t}>0} P_{s,t}\Delta S_{i,s,t}, \sum_{s|\Delta S_{i,s,t}\leq 0} P_{s,t}\left|\Delta S_{i,s,t}\right|\right)}{\frac{1}{2}\sum_{s} P_{s,t-1}S_{i,s,t-1} + \frac{1}{2}\sum_{s} P_{s,t}S_{i,s,t}}, \tag{49}$$

where $P_{s,t}$ is the price and $S_{i,s,t}$ is the number of shares of stock S held by institution i at the end of quarter t. The churn rate is equal to the minimum of net purchases and sales during quarter t as a fraction of the institution's average value over the two quarters, and it is used to measure the turnover of each institution's portfolio. Due to the presence of the minimum operator, institutions must both buy and sell large fractions of their portfolios to register a high churn rate. The mean churn rate is 0.12 and the standard deviation is 0.14, indicating a high degree of right skewness as the minimum churn rate is zero.

If institutions specialize in the rate at which they trade, then the churn rate should be persistent over time within institutions. Figure 5 plots the sample autocorrelations $\mathbf{corr}(c_{i,t}, c_{i,t-\Delta t})$ for $\Delta t \geq 2$. The churn rate strongly persists over time, with an autocorrelation of 0.51 over 10 years and 0.21 over 30 years.[19] We also find that institution fixed effects (δ_i) account for 65 percent of the variance in the churn rate in the regression $c_{i,t} = \delta_i + \varepsilon_{i,t}$, where $\varepsilon_{i,t}$ a residual.

4.3 Fund performance

To separate institutions according to their trading frequency, at each quarter t we divide institutions into deciles, denoted d, based on the mean of $c_{i,t}$ over the previous five years.[20] For simplicity, we restrict attention to top and bottom deciles ($d = 10$ and $d = 1$). Table 1 lists several institutions in these extreme deciles during the most recent quarter in our data. The top decile contains several well-known quantitative and high-frequency trading firms, whereas the bottom contains endowments and insurance companies.

Table 1: Institutions in the top and bottom deciles of churn rate in the fourth quarter of 2015

[19] Institution identifiers can be reassigned over time in the 13F data, leading to measurement error that biases the longer-term autocorrelations towards zero.

[20] We restrict attention to institutions for which the $t+1$ return on some of their holdings appears in CRSP, as these are the only institutions that can be analyzed in our return regressions.

Top decile	Bottom decile
Arrowstreet Capital	Berkshire Hathaway
Citadel	Bill & Melinda Gates Foundation
Dynamic Capital Management	Lilly Endowment
Ellington Management Group	Longview Asset Management
Quantlab	MetLife
Renaissance Technologies	New York State Teachers' Retirement System
Soros Fund Management	University of Notre Dame
Virtu Financial	University of Chicago

At the beginning of quarter t, we average the portfolio holdings of all the funds in each decile d at the end of quarter $t-1$ (with equal weight on each fund) and then track the returns on that aggregate decile-level portfolio over subsequent quarters, reinvesting proceeds from any delistings in the remaining stocks in the portfolio according to their value weights at that time. The return during quarter t of the decile d portfolio formed in quarter $t-k$ is denoted $r_{d,k,t}$.

We measure the performance of each portfolio by its alpha,

$$r_{d,k,t} - r_t^f = \alpha_{d,k} + \beta_{d,k} F_t + \varepsilon_{d,k,t}, \tag{50}$$

F_t is a vector of market risk factors; $F_t = r_t^m - r_t^f$ in the CAPM specification and $F_t = (r_t^m - r_t^f, r_t^{smb}, r_t^{hml})'$ in the Fama-French specification (r_t^{smb} and r_t^{hml} are returns on the SMB and HML portfolios, respectively). We focus on returns over the first two years after portfolio formation by estimating (50) only for $k \leq 8$.

Figure 6 displays alphas in the two specifications. For simplicity, we compare the alphas in the first quarter ($\alpha_{d,1}$) to those in the following seven quarters, $\alpha_{d,>1} \equiv \frac{1}{7}\sum_{j=2}^{8} \alpha_{d,j}$. The holdings of high-turnover funds out-perform more during the first quarter, while those of low-turnover funds out-perform more during subsequent quarters. The difference in differences $(\alpha_{10,1} - \alpha_{10,>1}) - (\alpha_{1,1} - \alpha_{1,>1})$, which measures the relative out-performance of high churn holdings versus low churn holdings at short versus long horizons, is equal to 0.005 in both specifications and is significant at the 5 percent level.[21] So, consistent with the model, high-turnover funds hold stocks that outperform relatively more in the short-run, while low-turnover funds hold assets that display more persistent outperformance.

5 The effects of eliminating high-frequency trade

Recently there has been interest in policies that might restrict high-frequency trading. Some of those policies are aimed at investors who trade at the very highest frequencies (such as the CFTC's recently proposed Regulation AT; see CFTC (2016)). But there are also proposals to discourage

[21] Yan and Zhang (2009) similarly find that the fraction of the outstanding shares of a stock held by high-churn institutions predicts subsequent returns, while the fraction held by low-churn institutions does not.

even portfolio turnover at the monthly or annual level.[22] There are two ways to interpret such policies. One would be that regulators might impose a tax on trading, which would simply represent a transaction cost. Such a regulation would obviously have the strongest effects on high-frequency traders (that result can be derived in an extension to the present setting with trading costs), but it would ultimately affect all trading strategies. A more targeted policy would be one that simply outlawed following a trading strategy in which positions fluctuate at frequencies above some bound. Such a policy is straightforward to analyze in our framework.

We show in this section that a policy that restricts high-frequency trading by professional investors (as opposed to liquidity traders) reduces liquidity and price informativeness and increases return volatility at high frequencies. At the frequencies not targeted by the policy, though, price informativeness is, if anything, increased.

5.1 The policy

The specific policy we analyze is a prohibition against any trading by sophisticated investors at high frequencies. In particular, there exists a $j_p \in (0, \frac{T}{2}]$ such that each sophisticated investor i is forced to choose $q_{i,j} = 0$ for $j \geq j_p$. The policy allows some trade by sophisticated investors because $j_p > 0$; at the very least, each trader i can choose a non-zero value of the low-frequency holding $q_{i,0}$.

No equilibrium exists in our main model with the addition of this policy because liquidity demand is perfectly inelastic, meaning that the markets at the restricted frequencies $j \geq j_p$ cannot clear. We therefore consider a simple extension to the baseline model where liquidity demand is somewhat elastic with respect to prices:

$$Z_t = \tilde{Z}_t + kP_t \tag{51}$$

$$z_j = \tilde{z}_j + kp_j, \tag{52}$$

where $\tilde{Z}_t$ is the exogenous supply process and $k > 0$ is the response of supply to prices. Appendix H solves this extended version of the model and obtains a water-filling equilibrium similar to that in solution 2, except that the λ_j functions now depend on k. In this section, λ_j denotes the function relevant for the equilibrium of the model with $k > 0$.

5.2 The reallocation of attention

The policy can change the total amount of investor attention, $f_{avg,j}^{-1}$, allocated to each frequency j. For frequencies $j \geq j_p$ that are regulated, the marginal benefit to investor i of allocating attention to j now equals 0 because $q_{i,j} = 0$. As a result, $f_{i,j}^{-1} = 0$ for all i and $j \geq j_p$, so the total investor

[22]The US tax code, for example, encourages holding assets for at least a year through the higher tax rates on short-term capital gains. There have been recent proposals to further expand such policies (a plan to create a schedule of capital gains tax rates that declines over a period of six years was attributed to Hillary Clinton during the 2016 US Presidential election; see Auxier et al. (2016)).

attention to each regulated frequency equals 0. If investors were allocating any attention to these frequencies before the policy, then aggregate investor attention on these high frequencies falls.

Interestingly, total investor attention on low frequencies may *increase* as a result of the policy. Recall the two heretofore equivalent formulations of the investor attention problem: (26), which assumes a fixed endowment of attention to be allocated across frequencies, and (28), which assumes a fixed marginal cost of attention. Under (26), a reduction in attention to restricted frequencies *increases* attention to the unrestricted frequencies $j < j_p$, as the marginal benefit of spending attention on these frequencies is always positive. Under (28), a reduction in attention to the restricted frequencies has no effect on attention to other frequencies, as the policy does not effect the marginal benefits and costs of attention at those frequencies. The following lemma, which is proved in an appendix, formally states these results.

Lemma 5 *If without the policy $f_{avg,j}^{-1} = 0$ for all $j \geq j_p$, then the policy leaves $f_{avg,j}^{-1}$ unaffected for all j. If without the policy $f_{avg,j}^{-1} > 0$ for some $j \geq j_p$, then the policy:*

- *lowers $f_{avg,j}^{-1}$ to 0 for all $j \geq j_p$;*
- *weakly increases $f_{avg,j}^{-1}$ for all $j < j_p$, strictly for some and for any $f_{avg,j}^{-1} > 0$ under* (26)*; and*
- *leaves $f_{avg,j}^{-1}$ unaffected for all $j < j_p$ under* (28).

The formulation in (26) fits a world in which there is some finance-specific attention or "skill." In this world, high-frequency traders substitute to become low-frequency traders when high-frequency trade is made illegal. Kierkoben, Leuven, and Mogstad (2017) present evidence that university students have comparative advantage in the fields of study (such as business, engineering, and medicine) that they choose. If such comparative advantage exists for finance generally, then (26) may fit the industry better than (28).

5.3 Price informativeness and return volatility

Presumably, the goal of regulating high-frequency trade is to improve welfare. Financial markets affect welfare through two channels we may analyze within our model. First, welfare is higher when prices convey more information about underlying cash flows, as argued by Hayek (1945). In our model, price informativeness at frequency j equals $Var(d_j \mid p)$, where the variance is taken over all possible realizations of dividends d and liquidity demand z. Second, as long as individuals dislike risk, welfare is higher when the volatility of returns is lower while expected cash flows are held constant. In our model, the volatility of returns at frequency j equals $Var(d_j - p_j)$, where the variance is again taken over d and z. We now explain how the policy affects welfare according to these two measures.

The following result expresses price informativeness in terms of the total attention $f_{avg,j}^{-1}$ allocated by sophisticated investors to each frequency j:

Result 6 *The informativeness of prices about dividends at frequency j equals*

$$Var(d_j \mid p)^{-1} = \left(\rho f_{avg,j}^{-1}\right)^2 f_{Z,j}^{-1} + f_{D,j}^{-1}. \tag{53}$$

Intuitively, price informativeness rises when the total attention $f_{avg,j}^{-1}$ rises.

Result 6 combines with Lemma 5 to offer predictions on the effect of the policy on price informativeness at various frequencies. At the regulated frequencies $j \geq j_p$, price informativeness weakly declines, strictly at any frequencies that were receiving attention before the policy. Without the price discovery provided by sophisticated investors, prices become completely uninformative at high frequencies. Mathematically, the lack of information holds because the conditional variance $Var(d_j \mid p)$ equals the prior variance $Var(d_j) = f_{D,j}$ when $f_{avg,j}^{-1} = 0$. Economically, at high frequencies prices reveal only liquidity demand Z, which is uninformative about dividends D.

This lack of information at high frequencies is counterbalanced by a possible increase in information about dividends at low frequencies. As shown by Lemma 5, attention to some low frequencies (for example, those that were already receiving attention) increases as a result of the policy under (26). Result 6 implies that the informativeness of prices at those frequencies grows. The extra attention from sophisticated investors moves prices closer to dividends, making p more informative about d_j.

The policy presents a trade-off: prices become (weakly) less informative about the high-frequency components of dividends, while they become (weakly) more informative about the low-frequency components of dividends. If policy-makers believe that low-frequency dividend information is more important for the economy, then the policy may improve welfare. As an example, suppose that only the sample mean of dividends is informative for economic decisions. This mean $T^{-1}\sum_{t=1}^{T} D_t$ is proportional to d_0, the lowest frequency dividend. Under the conditions given in Lemma 5, prices are more informative about this important piece of information with the policy that without it.[23]

A similar trade-off exists with the volatility of returns. The following result solves for this volatility with and without the policy:

Result 7 *Without the policy, return volatility equals* $Var(d_j - p_j) = \min\left(\bar{\lambda}, \lambda_j(0)\right)$. *With the policy, return volatility is given by*

$$Var(d_j - p_j) = \begin{cases} \min\left(\bar{\lambda}_p, \lambda_j(0)\right) & \text{if } j < j_p \\ f_{D,j} + f_{\tilde{Z},j} k^{-2} & \text{if } j \geq j_p, \end{cases} \tag{54}$$

where $\lambda_j(0) = f_{D,j} + f_{\tilde{Z},j}\left(k + \rho f_{D,j}^{-1}\right)^{-2}$. *Under* (28), $\bar{\lambda}_p = \bar{\lambda}$; *under* (26), $\bar{\lambda}_p \leq \bar{\lambda}$ *and is given by*

$$T^{-1} \sum_{j,j':\lambda_j(0)>\bar{\lambda}_p \wedge j,j'<j_p} \lambda_j^{-1}\left(\bar{\lambda}_p\right) = \bar{f}^{-1}. \tag{55}$$

[23] In appendix I, we study the informativeness of prices for more general moving averages of dividends.

The policy weakly increases return volatility at high frequencies—strictly at those frequencies where liquidity demand exists and $f_{\tilde{Z},j} > 0$. Volatility rises for two reasons. First, the frequencies that were receiving attention without the policy lose that attention, and this loss of attention moves prices further from dividends on average. Through this channel, volatility rises from $\bar{\lambda}$ to $\lambda_j(0)$. A second channel can push the volatility even higher than $\lambda_j(0)$, however. Investors can no longer trade at high frequencies, so they no longer absorb liquidity demand, which investors would do even if they were not allowed to pay attention to dividends at those frequencies. This inability to buffer liquidity demand pushes volatility from $\lambda_j(0)$ to $f_{D,j} + f_{\tilde{Z},j}k^{-2}$, which is higher than $\lambda_j(0)$ as long as liquidity demand exists at frequency j. This second channel is stronger when the risk-bearing capacity ρ of investors is higher.

Return volatility declines at some low frequencies if $\bar{\lambda}_p$ is less than $\bar{\lambda}$. As can be seen from result 7, $\bar{\lambda}_p < \bar{\lambda}$ when some attention is allocated to high frequencies ($\lambda_j(0) > \bar{\lambda}$ for some $j \geq j_p$) and (26) holds. In this case, investors reallocate some of their fixed stock of attention from high frequencies to low frequencies, and the shadow value of attention falls from $\bar{\lambda}$ to $\bar{\lambda}_p$. At low frequencies receiving attention, return volatility falls as a consequence of the increased attention. If $\bar{\lambda} = \bar{\lambda}_p$, then return volatility stays constant at low frequencies.

We examine the policy's quantitative implications for attention and return volatility in the context of the calibration used above, which assumes (26). The top panel of figure 3 plots $f_{avg,j}^{-1}$ both without and with a policy prohibiting trade at frequencies above $\omega = 3$ (cycle lengths shorter than 2.1 periods). As shown in the figure, no information is acquired at the regulated frequencies. Investors allocate their attention elsewhere, so information acquisition increases at other frequencies. The bottom panel plots return volatility with and without the policy, as well as in a hypothetical in which investors are restricted from paying attention to but not from trading at high frequencies; in this case, the variance equals $\min\left(\bar{\lambda}_p, \lambda_j(0)\right)$. Return volatility weakly declines at unrestricted frequencies, and it increases at restricted frequencies, with most of the policy's effect coming from the prohibition against trading and not just from the endogenous response of less attention.

In this calibrated example, the policy sharply increases the volatility of high-frequency returns primarily by prohibiting investors from absorbing high-frequency variation in liquidity demand, while slightly lowering the volatility of low-frequency returns by inducing a reallocation of investor attention to low-frequency dividend information. These changes may increase or decrease welfare depending on the relative cost of volatility at different horizons for participants in the market.

6 Conclusion

This paper develops a model in which there are many different investors who all trade at different frequencies. Investors in real-world markets follow countless strategies that are associated with rates of turnover that differ by multiple orders of magnitude. We show that in fact it is entirely natural that investors would differentiate along the dimension of investment frequency.

It has been standard in the literature for decades to focus on factors or principal components when studying the cross-section of asset returns. For stationary time series, the analog to factors or principal components is the set of fluctuations at different frequencies. So just as it seems natural for investors to focus on particular factors in the cross-section of returns, e.g. value stocks, a particular industry, or a particular commodity, it is also natural for investors to focus on fluctuations in fundamentals at a particular frequency, like quarters, business cycles, or decades.

Such an attention allocation problem can be solved using a combination of standard tools from time series econometrics and the literature on equilibria in financial markets. We show that the model fits a wide range of basic stylized facts about financial markets: investors can be distinguished by turnover rates; trading frequencies align with research frequencies; volume is driven primarily by high-frequency traders; and the positions of informed traders forecast returns at a horizon similar to their holding period.

Since the model has a rigorous concept of what being a high- or low-frequency trader entails, it is particularly useful for studying the effects of regulatory policies that would restrict trade at certain frequencies, whether by outlawing it or by simply making it more costly. We find that not only do such policies reduce the informativeness of prices at those frequencies, they also reduce liquidity and increase return volatility. In fact, return volatility will in general be raised even above where it would be in the complete absence of information, since eliminating active traders from the market removes their risk-bearing capacity. Because the allocation of attention to high- and low-frequency trading is endogenous, return volatility may decrease at lower frequencies as a result of such policies.

At this point, the primary drawback of the model in our view is that it is not fully dynamic. In a certain sense we have to assume that investors do not update information sets over time. While that simplification does not interfere with the model's ability to match a wider range of basic facts about financial markets, a simple desire for realism suggests that incorporating dynamic learning is an obvious next step.

References

Admati, Anat R, "A Noisy Rational Expectations Equilibrium for Multi-Asset Securities Markets," *Econometrica*, 1985, pp. 629–657.

Amihud, Yakov and Haim Mendelson, "Asset Pricing and the Bid-Ask Spread," *Journal of Financial Economics*, 1986, *17* (2), 223–249.

Auxier, Richard, Len Burman, Jim Nunns, Ben Page, and Jeff Rohaly, "An Updated Analysis of Hillary Clinton's Tax Proposals," October 2016.

Banerjee, Snehal, "Learning from Prices and the Dispersion in Beliefs," *Review of Financial Studies*, 2011, *24* (9), 3025–3068.

——— **and Brett Green**, "Signal or noise? Uncertainty and learning about whether other traders are informed," *Journal of Financial Economics*, 2015, *117* (2), 398–423.

——— **and Ilan Kremer**, "Disagreement and learning: Dynamic patterns of trade," *The Journal of Finance*, 2010, *65* (4), 1269–1302.

Basak, Suleyman, "Asset pricing with heterogeneous beliefs," *Journal of Banking & Finance*, 2005, *29* (11), 2849–2881.

Bhattacharya, Sudipto and Paul Pfleiderer, "Delegated Portfolio Management," *Journal of Economic Theory*, 1985, *36* (1), 1–25.

Brillinger, David R., *Time Series: Data Analysis and Theory*, McGraw Hill, 1981.

Brockwell, Peter J. and Richard A. Davis, *Time Series: Theory and Methods*, Springer, 1991.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan, "High-Frequency Trading and Price Discovery," *Review of Financial Studies*, 2014, *27* (8), 2267–2306.

Carhart, Mark M., "On Persistence in Mutual Fund Performance," *Journal of Finance*, 1997, *52* (1), 57–82.

Chen, Hsiu-Lang, Narasimhan Jegadeesh, and Russ Wermers, "The value of active mutual fund management: An examination of the stockholdings and trades of fund managers," *Journal of Financial and quantitative Analysis*, 2000, *35* (03), 343–368.

Commission, Commodity Futures Trading, "Regulation Automated Trading; Supplemental notice of proposed rulemaking," in "Federal Register" 2016.

Cuny, Charles J., "Why Derivatives on Derivatives? The Case of Spread Futures," *Journal of Financial Intermediation*, 2006, *15*, 132–159.

Dávila, Eduardo and Cecilia Parlatore, "Trading Costs and Informational Efficiency," 2016. Working paper.

DeFusco, Anthony A., Charles G. Nathanson, and Eric Zwick, "Speculative Dynamics of Prices and Volume," 2016. Working paper.

Dew-Becker, Ian and Stefano Giglio, "Asset Pricing in the Frequency Domain: Theory and Empirics," *Review of Financial Studies*, 2016. Forthcoming.

Diamond, Douglas W and Robert E Verrecchia, "Information aggregation in a noisy rational expectations economy," *Journal of Financial Economics*, 1981, *9* (3), 221–235.

Gompers, Paul A and Andrew Metrick, "Institutional Investors and Equity Prices," *Quarterly Journal of Economics*, 2001, pp. 229–259.

Griffin, John M and Jin Xu, "How smart are the smart guys? A unique view from hedge fund stock holdings," *Review of Financial Studies*, 2009, *22* (7), 2531–2570.

Grossman, Sanford J. and Joseph Stiglitz, "On the Impossibility of InfoInformation Efficient Markets," *American Economic Review*, 1980, *70*, 393–408.

Hayek, F. A., "The Use of Knowledge in Society," *American Economic Review*, 1945, *35* (4), 519–530.

He, Hua and Jiang Wang, "Differential information and dynamic behavior of stock trading volume," *Review of Financial Studies*, 1995, *8* (4), 919–972.

Hellwig, Martin F, "On the aggregation of information in competitive markets," *Journal of Economic Theory*, 1980, *22* (3), 477–498.

Hong, Harrison and Jeremy C Stein, "Disagreement and the stock market," *The Journal of Economic Perspectives*, 2007, *21* (2), 109–128.

Hopenhayn, Hugo A and Ingrid M Werner, "Information, Liquidity, and Asset Trading in a Random Matching Game," *Journal of Economic Theory*, 1996, *68* (2), 349–379.

Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp, "Rational attention allocation over the business cycle," Technical Report, National Bureau of Economic Research 2016.

Kirkeboen, Lars, Edwin Leuven, and Magne Mogstad, "Field of Study, Earnings, and Self-Selection," *Quarterly Journal of Economics*, 2017. Forthcoming.

Nagel, Stefan, "Short sales, institutional investors and the cross-section of stock returns," *Journal of Financial Economics*, 2005, *78* (2), 277–309.

Scheinkman, José A. and Wei Xiong, "Overconfidence and Speculative Bubbles," *Journal of Political Economy*, 2003, *111* (6), 1183–1220.

Shao, Xiaofeng and Wei Biao Wu, "Asymptotic Spectral Theory for Nonlinear Time Series," *The Annals of Statistics*, 2007, *35* (4), 1773–1801.

Shumway, Robert H. and David T. Stoffer, *Time Series Analysis and Its Applications*, New York: Springer, 2011.

Spiegel, Matthew, "Stock price volatility in a multiple security overlapping generations model," *Review of Financial Studies*, 1998, *11* (2), 419–447.

Stoughton, Neal M, "Moral Hazard and the Portfolio Management Problem," *The Journal of Finance*, 1993, *48* (5), 2009–2028.

Townsend, Robert M, "Forecasting the forecasts of others," *The Journal of Political Economy*, 1983, pp. 546–588.

Turley, Robert, "Informative Prices and the Cost of Capital Markets," 2012. Working paper.

Wang, Jiang, "A model of intertemporal asset prices under asymmetric information," *The Review of Economic Studies*, 1993, *60* (2), 249–282.

——, "A model of competitive stock trading volume," *Journal of political Economy*, 1994, pp. 127–168.

Watanabe, Masahiro, "Price volatility and investor behavior in an overlapping generations model with information asymmetry," *The Journal of Finance*, 2008, *63* (1), 229–272.

Weller, Brian, "Efficient Prices at Any Cost: Does Algorithmic Trading Deter Information Acquisition?," 2016. Working paper.

Yan, Xuemin Sterling and Zhe Zhang, "Institutional Investors and Equity Returns: Are Short-Term Institutions Better Informed?," *Review of Financial Studies*, 2009, *22* (2), 893–924.

A Non-stationary fundamentals

If fundamentals are non-stationary, e.g. if D_t has a unit root, then Σ_D is no longer Toeplitz and our results do not hold. In that case, we assume that D_0 is known by all agents and that the distribution of $\Delta D_t \equiv D_t - D_{t-1}$ is known, with covariance matrix $\Sigma_{\Delta D}$. Then the entire problem can simply be rescaled by defining $\tilde{P}_t \equiv P_t - D_{t-1}$, so that

$$\begin{aligned} R_t &= D_t - P_t && (56) \\ &= \Delta D_t - \tilde{P}_t && (57) \end{aligned}$$

Our analysis then applies to $\tilde{P}_t$ and ΔD_t, with $Q_{i,t}$ continuing to represent the number of forward contracts on D_t that agent i buys. That is, we are allowing agents to condition demand $Q_{i,t}$ not just on signals and prices, but also the level of D_{t-1}, simply through differencing.

B Proof of lemma 1

Gray (2006) shows that for any circulant matrix (a matrix where row n is equal to row $n-1$ circularly shifted right by one column, and thus one that is uniquely determined by its top row), the discrete Fourier basis, $u_j = [\exp(i\omega_j t), t = 0, ..., T-1]'$ for $j \in \{0, ..., T-1\}$, is the set of eigenvectors.

Let Σ be a symmetric Toeplitz matrix with top row $[\sigma_0, \sigma_1, ..., \sigma_{T-1}]$. Define the function $circ(x)$ to be a circulant matrix with σ_{circ} as its top row. Define a vector σ

$$\sigma \equiv [\sigma_0, \sigma_1 + \sigma_{T-1}, \sigma_2 + \sigma_{T-2}, ..., \sigma_{T-2} + \sigma_2, \sigma_{T-1} + \sigma_1]' \tag{58}$$

Following Rao (2016), we "approximate" Σ by the circulant matrix $\Sigma_{circ} \equiv circ(\sigma_{circ})$. Since Σ_{circ} is symmetrical, one may observe that its eigenvalues repeat in the sense that $u_j' \Sigma_{circ} = u_{T-j}' \Sigma_{circ}$ for $0 < j < T$. Since pairs of eigenvectors with matched eigenvalues can be linearly combined to yield alternative eigenvectors, it immediately follows that the matrix Λ from the main text contains a full set of eigenvectors for Σ_{circ}. The associated eigenvalues are

$$f_{\Sigma_{circ}}(\omega_j) = \sigma_0 + 2\sum_{t=1}^{T-1} \sigma_t cos(\omega_j t) \tag{59}$$

We can write this relationship more compactly as:

$$\begin{aligned} \Sigma_{circ}\Lambda &= \Lambda f_\Sigma \\ \Lambda' \Sigma_{circ} \Lambda &= f_\Sigma \end{aligned}$$

where the $T \times T$ diagonal matrix f_Σ is given by:

$$f_\Sigma = diag\left(f_\Sigma(\omega_0), f_\Sigma(\omega_1), ..., f_\Sigma\left(\omega_{\frac{T}{2}}\right), f_\Sigma(\omega_1), f_\Sigma(\omega_2), ..., f_\Sigma\left(\omega_{\frac{T}{2}-1}\right)\right)'.$$

The approximate diagonalization of the matrix Σ consists in writing:

$$\begin{aligned}\Lambda'\Sigma\Lambda &= f_\Sigma + R_\Sigma \\ \text{where } R_\Sigma &\equiv \Lambda'(\Sigma - \Sigma_{circ})\Lambda\end{aligned}$$

By direct inspection of the elements of $\Sigma - \Sigma_{circ}$, one may see that the m, n element of R_Σ, denoted $R_\Sigma^{m,n}$ satisfies (defining λ_m to be the mth column of Λ and $\lambda_{m,n}$ to be its m, n element)

$$\begin{aligned}R_\Sigma^{m,n} &\equiv \lambda_m'(\Sigma - \Sigma_{circ})\lambda_n && (60)\\ &= \sum_{i=1}^{T}\sum_{j=1}^{T}\lambda_{m,i}\lambda_{n,j}(\Sigma - \Sigma_{circ})^{m,n} && (61)\\ &\leq \sum_{i=1}^{T}\sum_{j=1}^{T}\frac{2}{T}(\Sigma - \Sigma_{circ})^{m,n} && (62)\\ &\leq \frac{4}{T}\sum_{j=1}^{T-1} j\,|\sigma_j| && (63)\end{aligned}$$

where $(\Sigma - \Sigma_{circ})^{m,n}$ is the m, n element of $(\Sigma - \Sigma_{circ})$. So R_Σ is bounded elementwise by a term of order T^{-1}. One may show that the weak norm satisfies $|\cdot| \leq \sqrt{T}\,|\cdot|_{\max}$, where $|\cdot|_{\max}$ denotes the elementwise max norm, which thus yields the result that $|\Lambda\Sigma\Lambda' - diag(f_\Sigma)| \leq bT^{-1/2}$ for some b.

B.1 Convergence in distribution and $\bar{O}$ bounds

Define the notation $\Rightarrow$ to mean that $\Lambda X \Rightarrow N\left(0, \hat{\Sigma}_X\right)$ if $\Lambda X \sim N(0, \Sigma_X)$ and $\left|\hat{\Sigma}_X - \Sigma_X\right| = \bar{O}\left(T^{-1/2}\right)$.

The notation $\bar{O}$ indicates

$$|A - B| = \bar{O}\left(T^{-1/2}\right) \Longleftrightarrow |A - B| \leq bT^{-1/2} \quad (64)$$

for some constant b and for all T. This is a stronger statement than typical big-O notation in that it holds for all T, as opposed to holding only for some sufficiently large T.

Trigonometric transforms of stationary time series converge in distribution under more general conditions. See Shumway and Stoffer (2011), Brillinger (1981), and Shao and Wu (2007).

C Derivation of solution 1

Since the optimization is entirely separable across frequencies (confirmed below), we can solve everything in scalar terms. To save notation, we suppress the j subscripts indicating frequencies in this section when they are not necessary for clarity. So in this section f_D, for example, is a scalar representing the spectral density of fundamentals at some arbitrary frequency.

C.1 Statistical inference

We guess that prices take the form

$$p = a_1 d - a_2 z \tag{65}$$

The joint distribution of fundamentals, signals, and prices is then

$$\begin{bmatrix} d \\ y_i \\ p \end{bmatrix} \sim N\left(0, \begin{bmatrix} f_D & f_D & a_1 f_D \\ f_D & f_D + f_i & a_1 f_D \\ a_1 f_D & a_1 f_D & a_1^2 f_D + a_2^2 f_Z \end{bmatrix}\right) \tag{66}$$

The expectation of fundamentals conditional on the signal and price is

$$E\left[d \mid y_i, p\right] = \begin{bmatrix} f_D & a_1 f_D \end{bmatrix} \begin{bmatrix} f_D + f_i & a_1 f_D \\ a_1 f_D & a_1^2 f_D + a_2^2 f_Z \end{bmatrix}^{-1} \begin{bmatrix} y_i \\ p \end{bmatrix} \tag{67}$$

$$= [1, a_1] \begin{bmatrix} 1 + f_i f_D^{-1} & a_1 \\ a_1 & a_1^2 + a_2^2 f_Z f_D^{-1} \end{bmatrix}^{-1} \begin{bmatrix} y_i \\ p \end{bmatrix} \tag{68}$$

and the variance satisfies

$$\tau_i \equiv Var\left[d \mid y_i, p\right]^{-1} = f_D^{-1} \left(1 - \begin{bmatrix} 1 & a_1 \end{bmatrix} \begin{bmatrix} 1 + f_i f_D^{-1} & a_1 \\ a_1 & a_1^2 + a_2^2 f_Z f_D^{-1} \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ a_1 \end{bmatrix}\right)^{-1} \tag{69}$$

$$= \frac{a_1^2}{a_2^2} f_Z^{-1} + f_i^{-1} + f_D^{-1} \tag{70}$$

We use the notation τ to denote a posterior precision, while f^{-1} denotes a prior precision of one of the basic variables of the model. The above then implies that

$$E\left[d \mid y_i, p\right] = \tau_i^{-1} \left(f_i^{-1} y_i + \frac{a_1}{a_2^2} f_Z^{-1} p\right) \tag{71}$$

C.2 Demand and equilibrium

The agent's utility function is (where variables without subscripts here indicate vectors),

$$U_i = \max_{\{Q_{i,t}\}} \rho^{-1} E_{0,i}\left[T^{-1} Q_i' (D-P)\right] - \rho^{-2} Var_{0,i}\left[T^{-1/2} Q_i' (D-P)\right] \tag{72}$$

$$= \max_{\{Q_{i,t}\}} \rho^{-1} E_{0,i}\left[T^{-1} q_i' (d-p)\right] - \rho^{-2} Var_{0,i}\left[T^{-1/2} q_i' (d-p)\right] \tag{73}$$

$$= \max_{\{Q_{i,t}\}} \rho^{-1} T^{-1} \sum_{j=0}^{T-1} q_{i,j} E_{0,i}\left[(d_j - p_j)\right] - \rho^{-2} T^{-1} \sum_{j=0}^{T-1} q_{i,j}^2 Var_{0,i}\left[d_j - p_j\right] \tag{74}$$

where the last line follows by imposing the asymptotic independence of d across frequencies (we analyze the error induced by that approximation below). The utility function is thus entirely separable across frequencies, with the optimization problem for each $q_{i,j}$ independent from all others.

Taking the first-order condition associated with the last line above for a single frequency, we obtain

$$q_i = \rho \tau_i E\left[d - p \mid y_i, p\right]$$
$$= \rho_i \left(f_i^{-1} y_i + \left(\frac{a_1}{a_2^2} f_Z^{-1} - \tau_i \right) p \right)$$

Summing up all demands and inserting the guess for the price yields

$$z = \int_i \rho \left(f_i^{-1} y_i + \left(\frac{a_1}{a_2^2} f_Z^{-1} - \tau_i \right) (a_1 d - a_2 z) \right) di \tag{75}$$

$$= \int_i \rho \left(f_i^{-1} d + \left(\frac{a_1}{a_2^2} f_Z^{-1} - \tau_i \right) (a_1 d - a_2 z) \right) di \tag{76}$$

Where the second line uses the law of large numbers. Matching coefficients then yields

$$\int_i \rho \left(\frac{a_1}{a_2^2} f_Z^{-1} - \tau_i \right) di = -a_2^{-1} \tag{77}$$

$$\int_i \rho f_i^{-1} + \rho \left(\frac{a_1}{a_2^2} f_Z^{-1} - \tau_i \right) a_1 di = 0 \tag{78}$$

and therefore

$$\rho \int_i f_i^{-1} di = \frac{a_1}{a_2} \tag{79}$$

Inserting the expression for τ_i into (77) yields

$$a_1 = \frac{\frac{a_1}{a_2} + \rho \left(\frac{a_1}{a_2}\right)^2 f_Z^{-1}}{\rho \left(\int_i f_i^{-1} di + f_D^{-1} + \left(\frac{a_1}{a_2}\right)^2 f_Z^{-1} \right)} \tag{80}$$

Now define aggregate precision to be

$$f_{avg}^{-1} \equiv \int_i f_i^{-1} di \tag{81}$$

We then have

$$\begin{aligned} \tau_i &= \frac{a_1^2}{a_2^2} f_Z^{-1} + f_i^{-1} + f_D^{-1} & (82) \\ \tau_{avg} &\equiv \int \tau_i di = \left(\rho f_{avg}^{-1}\right)^2 f_Z^{-1} + f_{avg}^{-1} + f_D^{-1} & (83) \end{aligned}$$

$$\begin{aligned} a_1 &= \tau_{avg}^{-1} \left(f_{avg}^{-1} + \left(\rho f_{avg}^{-1}\right)^2 f_Z^{-1} \right) = 1 - \frac{f_D^{-1}}{\tau_{avg}} = \frac{\tau_{avg} - f_D^{-1}}{\tau_{avg}} & (84) \\ a_2 &= \frac{a_1}{\rho f_{avg}^{-1}} & (85) \end{aligned}$$

C.3 Proof of proposition 1

In the time domain, the solution from Admati (1985) is

$$\begin{aligned} P &= A_1 D - A_2 Z & (86) \\ A_1 &\equiv I - S_{avg}^{-1} \Sigma_D^{-1} & (87) \\ A_2 &\equiv \rho^{-1} A_1 \Sigma_{avg} & (88) \end{aligned}$$

Standard properties of norms yield the following result. If $|A - B| = \bar{O}\left(T^{-1/2}\right)$ and $|C - D| = \bar{O}\left(T^{-1/2}\right)$, then

$$\begin{aligned} |cA - cB| &= \bar{O}\left(T^{-1/2}\right) & (89) \\ \left|A^{-1} - B^{-1}\right| &= \bar{O}\left(T^{-1/2}\right) & (90) \\ |(A + C) - (B + D)| &= \bar{O}\left(T^{-1/2}\right) & (91) \\ |AC - BD| &= \bar{O}\left(T^{-1/2}\right) & (92) \end{aligned}$$

In other words, convergence in weak norm carries through under addition, multiplication, and inversion. Since A_1 is a function of Toeplitz matrices using those operations, it follows that $|\Lambda' A_1 \Lambda - diag\,(a_1)| = \bar{O}\left(T^{-1/2}\right)$, and the same holds for A_2.

For the variance of prices, we define

$$\begin{aligned} R_1 &= A_1 - \Lambda diag\,(a_1)\,\Lambda' & (93) \\ R_2 &= A_2 - \Lambda diag\,(a_2)\,\Lambda' & (94) \end{aligned}$$

$$
\begin{aligned}
|Var[P-\Lambda p]| &\leq \left|R_1\Sigma_D R_1'\right| + \left|R_2\Sigma_Z R_2'\right| && (95)\\
&\leq |R_1\Sigma_D|\,|R_1| + |R_2\Sigma_Z|\,|R_2| && (96)\\
&\leq \|\Sigma_D\|\,|R_1|^2 + \|\Sigma_Z\|\,|R_2|^2 && (97)\\
&\leq K\left(|R_1|^2 + |R_2|^2\right) && (98)
\end{aligned}
$$

The first line follows from the triangle inequality; the second line comes from the sub-multiplicativity of the weak norm; the third line uses the fact that, as indicated by Gray (2006), for any two square matrices G, H, $||GH||_2 \leq \|G\|\,|H|$; and the last line follows from the assumption that the eigenvalues of Σ_D and Σ_Z are bounded by some K.

Since the weak norm is invariant under unitary transformations,

$$
|R_1| = \left|\Lambda' R_1 \Lambda\right| = \left|\Lambda' A_i \Lambda - diag\,(a_1)\right| \quad , \quad i = 1, 2.
$$

Therefore,

$$
\begin{aligned}
|Var[P-\Lambda P]| &\leq K\left(\left|\Lambda' A_1\Lambda - diag\,(a_1)\right|^2 + \left|\Lambda' A_2 \Lambda - diag\,(a_2)\right|^2\right) && (99)\\
&= \bar{O}\left(\frac{1}{T}\right) && (100)
\end{aligned}
$$

Since $\|\cdot\| \leq \sqrt{T}\,|\cdot|$, $\|Var[P^c - P]\| = \bar{O}\left(T^{-1/2}\right)$.

D Proof of lemma 2

Inserting the optimal value of $q_{i,j}$ into the utility function, we obtain

$$
E_{-1}[U_{i,0}] \equiv \frac{1}{2} E\left[T^{-1}\sum_{j=0}^{T-1}\tau_{i,j} E\left[d_j - p_j \mid y_{i,j}, p_j\right]^2\right] \tag{101}
$$

$U_{i,0}$ is utility conditional on an observed set of signals and prices. $E_{-1}[U_{i,0}]$ is then the expectation taken over the distributions of prices and signals.

$Var\left[E\left[d_j - p_j \mid y_{i,j}, p_j\right]\right]$ is the variance of the part of the return on portfolio j explained by $y_{i,j}$ and p_j, while $\tau_{i,j}^{-1}$ is the residual variance. The law of total variance says

$$
Var[d_j - p_j] = Var\left[E\left[d_j - p_j \mid y_{i,j}, p_j\right]\right] + E\left[Var\left[d_j - p_j \mid y_{i,j}, p_j\right]\right] \tag{102}
$$

where the second term on the right-hand side is just $\tau_{i,j}^{-1}$ and the first term is $E\left[E\left[d_j - p_j \mid y_{i,j}, p_j\right]^2\right]$ since everything has zero mean. The unconditional variance of returns is

$$
Var[d_j - p_j] = (1 - a_{1,j})^2 f_{D,j} + \frac{a_{1,j}^2}{\left(\rho f_{avg,j}^{-1}\right)^2} f_{Z,j} \tag{103}
$$

So then

$$E_{-1}\left[U_{i,0}\right]=\frac{1}{2}T^{-1}\sum_{j=0}^{T-1}\left(\left(1-a_{1,j}\right)^{2}f_{D,j}+\frac{a_{1,j}^{2}}{\left(\rho f_{avg}^{-1}\right)^{2}}f_{Z,j}\right)\tau_{i,j}-\frac{1}{2} \tag{104}$$

We thus obtain the result that agent i's expected utility is linear in the precision of the signals that they receive (since $\tau_{i,j}$ is linear in $f_{i,j}^{-1}$).

Furthermore,

$$\begin{aligned}\left(1-a_{1,j}\right)^{2}f_{D,j}+\frac{a_{1,j}^{2}}{\left(\rho f_{avg,j}^{-1}\right)^{2}}f_{Z,j} &= \tau_{avg,j}^{-2}f_{D,j}^{-1}+\rho^{-2}f_{avg,j}^{2}\tau_{avg,j}^{-2}\left(\tau_{avg,j}-f_{D,j}^{-1}\right)^{2}f_{Z,j} && (105)\\ &= \tau_{avg,j}^{-2}f_{D,j}^{-1}+\rho^{-2}\tau_{avg,j}^{-2}\left(\rho^{2}f_{avg}^{-1}f_{Z}^{-1}+1\right)^{2}f_{Z,j} && (106)\\ &= \tau_{avg,j}^{-2}\left(f_{D,j}^{-1}+\rho^{-2}\left(\rho^{2}f_{avg}^{-1}f_{Z}^{-1}+1\right)^{2}f_{Z,j}\right) && (107)\end{aligned}$$

So

$$E_{-1}\left[U_{i,0}\right]=\frac{1}{2}T^{-1}\sum_{j=0}^{T-1}\tau_{avg,j}^{-2}\left(f_{D,j}^{-1}+\rho^{-2}\left(\rho^{2}f_{avg}^{-1}f_{Z}^{-1}+1\right)^{2}f_{Z,j}\right)\left(\left(\rho f_{avg,j}^{-1}\right)^{2}f_{Z,j}^{-1}+f_{i,j}^{-1}+f_{D,j}^{-1}\right)-\frac{1}{2} \tag{108}$$

E Derivation of solution 2

Investors allocate attention, $f_{i,j}^{-1}$, to maximize $E_{-1}\left[U_{i,0}\right]$ subject to the constraint

$$\sum_{j,j'}f_{i,j}^{-1}\leq\bar{f}^{-1} \tag{109}$$

and that $f_{i,j}^{-1}=f_{i,j'}^{-1}$. Since the investors are maximizing a linear objective subject to a linear constraint, the optimal policy is clearly to allocate attention $f_{i,j}^{-1}$ only to the frequencies j at which the marginal benefit is equal to the maximum available marginal benefit.

Define the function λ_j

$$\lambda_{j}\left(x\right)\equiv\left(\left(\rho x\right)^{2}f_{Z,j}^{-1}+x+f_{D,j}^{-1}\right)^{-2}\left(f_{D,j}^{-1}+\rho^{-2}\left(\rho^{2}xf_{Z,j}^{-1}+1\right)^{2}f_{Z,j}\right)$$

then $\lambda_j\left(f_{avg,j}^{-1}\right)$ is the marginal benefit from attention to frequency j. Note that $d\lambda_j\left(x\right)/dx<0$. In equilibrium, then, there is a number $\bar{\lambda}$ such that

$$\lambda_{j}\left(f_{avg,j}^{-1}\right)\leq\bar{\lambda}\text{ for all }j \tag{110}$$

Now define $\mathcal{J}\left(\bar{\lambda}\right)$ to be the set of frequencies j such that $\lambda_j^{-1}\left(\bar{\lambda}\right)>0$.[24] That is the set of

[24] Technically, it is the set of frequencies for which $\lambda_j^{-1}\left(\min\left(\bar{\lambda},\lambda_j\left(0\right)\right)\right)>0$.

frequencies for which there is positive attention.

For any frequency that investors allocate attention to,

$$f_{avg,j}^{-1} = \lambda_j^{-1}\left(\bar{\lambda}\right) \tag{111}$$

$$f_{avg,j}^{-1} = \int f_{i,j}^{-1} di \tag{112}$$

Now

$$\sum_{j,j'\in\mathcal{J}} \int f_{i,j}^{-1} di = \int \sum_{j,j'\in\mathcal{J}} f_{i,j}^{-1} di \tag{113}$$

$$= \int \bar{f}^{-1} di = \bar{f}^{-1} \tag{114}$$

So then

$$\sum_{j,j'\in\mathcal{J}(\bar{\lambda})} \lambda_j^{-1}\left(\bar{\lambda}\right) = \sum_{j,j'\in\mathcal{J}(\bar{\lambda})} f_{avg,j}^{-1} = \bar{f}^{-1} \tag{115}$$

So $\bar{\lambda}$ is obtained by solving $\sum_{j,j'\in\mathcal{J}(\bar{\lambda})} \lambda_j^{-1}\left(\bar{\lambda}\right) = \bar{f}^{-1}$.

F Time horizon and investment

At first glance, the assumption of mean-variance utility over cumulative returns over a long period of time ($T \to \infty$) may appear to give investors an incentive to primarily worry about long-horizon performance, whereas a small value of T would make investors more concerned about short-term performance. In the present setting, that intuition is not correct – the $T \to \infty$ limit determines how detailed investment strategies may be, rather than incentivizing certain types of strategies.

The easiest way to see why the time horizon controls only the detail of the investment strategies is to consider settings in which T is a power of 2. If $T = 2^k$, then the set of fundamental frequencies is

$$\left\{2\pi j/2^k\right\}_{j=0}^{2^{k-1}} \tag{116}$$

For $T = 2^{k-1}$, the set of frequencies is

$$\left\{2\pi j/2^{k-1}\right\}_{j=0}^{2^{k-2}} = \left\{2\pi\left(2j\right)/2^k\right\}_{j=0}^{2^{k-2}} \tag{117}$$

That is, when T falls from 2^k to 2^{k-1}, the effect is to simply eliminate alternate frequencies. Changing T does not change the lowest or highest available frequencies (which are always 0 and π, respectively). It just discretizes the $[0,\pi]$ interval more coarsely; or, equivalently, it means that the matrix Λ is constructed from a smaller set of basis vectors.

When T is smaller – there are fewer available basis functions – Q and its frequency domain analog $q \equiv \Lambda' Q$ have fewer degrees of freedom and hence must be less detailed. So the effect of a small value of T is to make it more difficult for an investor to *isolate* particularly high- or low-frequency fluctuations in fundamentals (or any other narrow frequency range). But in no way does T cause the investor's portfolio to depend more on one set of frequencies than another. While we take $T \to \infty$, we will see that the model's separating equilibrium features investors who trade at both arbitrarily low and high frequencies, and T has no effect on the distribution of investors across frequencies.

G Specialization model

G.1 Existence and unicity of equilibrium with fixed costs

The proof below proceeds under the assumption that:

$$\bar{f}^{-1} - \kappa \geq \sum_{j,j'} \lambda_j^{-1}(\underline{\lambda}(0)),$$

where:

$$\underline{\lambda}(0) = \min_j \left(\lambda_j(0)\right),$$

so that all frequencies receive some attention in the aggregate allocation of attention. This is also without loss of generality and only to simplify notation.

We first prove that given the aggregate allocation of attention described in lemma 3, specialization is individually optimal. This attention allocation is given by:

$$\begin{aligned} f_{avg,j}^{-1} &= \gamma_j(\bar{f}^{-1} - \kappa) \\ &= \lambda_j^{-1}(\bar{\lambda}) \end{aligned} \tag{118}$$

The individual problem of an agent is then:

$$\begin{aligned} U_{-1,i}\left(f_{avg}^{-1}\right) &= \mathcal{C}\left(f_{avg}^{-1}\right) + \quad \frac{1}{2T} \max_{f_i^{-1}} \quad \bar{\lambda} \sum_{j,j'} f_{i,j}^{-1} \\ \text{s.t.} \qquad & f_{i,j}^{-1} \geq 0 \\ & tr(f_i^{-1}) + \kappa \textstyle\sum_{j,j'} \mathbb{1}_{\left\{f_{i,j}^{-1} > 0\right\}} \leq \bar{f}^{-1} \end{aligned} \tag{119}$$

The solution of this problem is as follows. Maximized utility is:

$$U_{-1,i}\left(f_{avg}^{-1}\right) = \mathcal{C}\left(f_{avg}^{-1}\right) + \frac{1}{2T}\bar{\lambda}\left(\bar{f}^{-1} - \kappa\right),$$

and the individual agent is indifferent across the $\frac{T}{2}+1$ different specialized strategies:

$$f_{i,l}^{-1}=\begin{cases}\left(\bar{f}^{-1}-\kappa\right) & \text{if } l\in\left\{0,\frac{T}{2}\right\}\\ \frac{1}{2}\left(\bar{f}^{-1}-\kappa\right) & \text{if } l\notin\left\{0,\frac{T}{2}\right\}\end{cases}\quad,\quad f_{i,-l}=0\quad,\qquad l=0,...,T-1.$$

To see why these $\frac{T}{2}+1$ specialized attention allocations are optimal, consider the following strategy, which allocates attention to two distinct frequencies:

$$\tilde{f}_{i,0}^{-1}>0\quad,\quad \tilde{f}_{i,1}^{-1}>0,\qquad \tilde{f}_{i,0}^{-1}+\tilde{f}_{i,1}^{-1}+2\kappa=\bar{f}^{-1}.$$

Since $\tilde{f}_{i,0}^{-1}+\tilde{f}_{i,1}^{-1}=\bar{f}^{-1}-2\kappa$, this allocation yields utility:

$$\mathcal{C}\left(f_{avg}^{-1}\right)+\frac{1}{2T}\bar{\lambda}\left(\bar{f}^{-1}-2\kappa\right).$$

However, the allocation:

$$f_{i,0}^{-1}=\tilde{f}_{i,0}^{-1}+\tilde{f}_{i,1}^{-1}+\kappa\quad,\quad f_{i,1}^{-1}=0$$

is then also feasible. It yields utility:

$$\mathcal{C}\left(f_{avg}^{-1}\right)+\frac{1}{2T}\bar{\lambda}\left(\bar{f}^{-1}-\kappa\right)>\mathcal{C}\left(f_{avg}^{-1}\right)+\frac{1}{2T}\bar{\lambda}\left(\bar{f}^{-1}-2\kappa\right),$$

and so is strictly preferred. Thus, specialization is consistent with individual optimality.

We next prove that, for any $\kappa>0$, there are no equilibria in which individual agents learn about two or more frequencies — i.e. no equilibria without specialization. Let $\left\{f_{avg,j}^{-1}\right\}$ denote the aggregate attention allocation associated with a particular equilibrium. Let:

$$\hat{\lambda}=\max_j \lambda_j(f_{avg,j}^{-1}),$$

$$\mathcal{J}=\arg\max_j \lambda_j(f_{avg,j}^{-1}).$$

Then, a variational argument similar to the proof above can be used to establish the two following facts:

1. the optimal strategies of individual agents never involve learning about frequencies $j\notin\mathcal{J}$;

2. the $\frac{T}{2}$ strategies where individual agents specialize in one particular $l\in\mathcal{J}$ yield strictly higher utility than strategies involving learning about 2 or more frequencies in $\mathcal{J}$.

Thus, there can be no equilibrium where it is privately optimal to learn about more than one frequency.

G.2 Interpretation of the fixed cost in the time-series domain

The idea behind lemma 4 is to use a smooth approximation to the sum of step functions $\sum_{j=0}^{T-1} \mathbb{1}_{\{f_{i,j}^{-1}>0\}}$, and see whether we can interpret the approximated constraint in the time domain. The following lemma will be used:

Lemma 6 *Let $\alpha > 0$ be a strictly positive number. Define the function $g_\alpha : \mathcal{M}_T \to \mathbb{R}_+$, where $\mathcal{M}_T$ is a subset (defined below) of the real-valued squared matrices of size T, by:*

$$\forall M \in \mathcal{M}_T, \quad g_\alpha(M) = tr\left[\left(I - e^{-2\alpha M}\right)\left(I + e^{-2\alpha M}\right)^{-1}\right].$$

Then, if M is diagonal with weakly positive entries,

$$\lim_{\alpha\to\infty} g_\alpha(M) = \sum_{j=0}^{T-1} \mathbb{1}_{\{M_{i,j}>0\}}.$$

This is the matrix analog a sequence of logistic functions to approximate the Heaviside function, but for matrices. (The proof of this lemma can be established in that way). This only works for a certain set of matrices $\mathcal{M}_T$, which are the matrices such that the inverse $\left(I + \exp^{-2\alpha M}\right)^{-1}$ exists. Those are the matrices for which the series:

$$I - e^{-2\alpha M} + e^{-4\alpha M} - e^{-6\alpha M} + ...$$

is absolutely convergent for some matrix norm. In the particular case where M is diagonal with weakly positive entries, then

$$e^{-2\alpha k M_{i,i}} \leq 1 \quad \forall k \leq 1,$$

so that the series above converges. Since f_i^{-1} is diagonal, this lemma says that:

$$\lim_{\alpha\to+\infty} g_\alpha(f_i^{-1}) = \sum_{j=0}^{T-1} \mathbb{1}_{\{f_{i,j}^{-1}>0\}},$$

so that the fixed cost assumption can now be interpreted as the limit of constraints of the form:

$$tr(f_i^{-1}) + g_\alpha(f_i^{-1}) \leq \bar{f}^{-1}.$$

The question is then whether g_α has some simple representation in the time domain.

Lemma 7 *Let Σ_i be the variance-covariance matrix of signals. Assume that Σ_i^{-1} exists and is in $\mathcal{M}_T$, and define: $f_i^{-1} = \Lambda\Sigma_i^{-1}\Lambda'$. Then:*

$$g_\alpha(\Sigma_i^{-1}) = g_\alpha(f_i^{-1}).$$

Proof. First note that, by application of Sylvester's formula, for any $l \geq 1$:

$$e^{-2\alpha l \Sigma_i^{-1}} = \Lambda e^{-2\alpha l f_i^{-1}} \Lambda'.$$

Next, since $\Sigma_i^{-1} \in \mathcal{M}_T$,

$$(I + e^{-2\alpha \Sigma_i^{-1}})^{-1} = \sum_{l \geq 0} (-1)^l e^{-2\alpha l \Sigma_i^{-1}}.$$

Therefore:

$$\begin{aligned}
g_\alpha(\Sigma_i^{-1}) &= tr\left((I - e^{-2\alpha\Sigma_i^{-1}})(I + e^{-2\alpha\Sigma_i^{-1}})^{-1}\right) \\
&= tr\left((I - e^{-2\alpha\Sigma_i^{-1}}) \sum_{l\geq 0}(-1)^l e^{-2\alpha l \Sigma_i^{-1}}\right) \\
&= tr\left(\sum_{l\geq 0}(-1)^l e^{-2\alpha l \Sigma_i^{-1}} - \sum_{l\geq 0}(-1)^l e^{-2\alpha (l+1) \Sigma_i^{-1}}\right) \\
&= tr\left(\sum_{l\geq 0}(-1)^l e^{-2\alpha l \Sigma_i^{-1}} - \sum_{l\geq 1}(-1)^{l-1} e^{-2\alpha l \Sigma_i^{-1}}\right) \\
&= tr\left(I + 2\sum_{l\geq 1}(-1)^l e^{-2\alpha l \Sigma_i^{-1}}\right) \\
&= T + 2\sum_{l\geq 1}(-1)^l tr\left(e^{-2\alpha l \Sigma_i^{-1}}\right) \\
&= T + 2\sum_{l\geq 1}(-1)^l tr\left(\Lambda e^{-2\alpha l f_i^{-1}} \Lambda'\right) \\
&= T + 2\sum_{l\geq 1}(-1)^l tr\left(e^{-2\alpha l f_i^{-1}} \Lambda'\Lambda\right) \\
&= T + 2\sum_{l\geq 1}(-1)^l tr\left(e^{-2\alpha l f_i^{-1}}\right) \\
&= g_\alpha(f_i^{-1}),
\end{aligned} \tag{120}$$

where going from the penultimate to the last line just the same steps as going from the first to the sixth line, but in reverse (the rest of the steps use the linearity and cyclicality of the trace and the fact that Λ is orthogonal). ■

H Proofs of specialization model predictions

H.1 Results 2 and 3

$$q_i = \rho\left(f_i^{-1} y_i + \left(\frac{a_1}{a_2^2} f_Z^{-1} - \tau_i\right) p\right)$$

The coefficient on $\tilde{\varepsilon}_i$ is f_i^{-1}. Straightforward but tedious algebra confirms that the coefficient on d is

$$\rho\left(f_{avg}^{-1} - f_i^{-1}\right)(a_1 - 1)$$

The coefficient on z is

$$1 + \rho\left(f_i^{-1} - f_{avg}^{-1}\right) a_2$$

We thus have

$$q_i = \rho\left(f_{avg}^{-1} - f_i^{-1}\right)(a_1 - 1)\, d + \left(1 + \rho\left(f_i^{-1} - f_{avg}^{-1}\right)\right) a_2 z \tag{121}$$

Now note that

$$r = (1 - a_1)\, d + d_2 z \tag{122}$$

So then

$$q_i = \rho\left(f_i^{-1} - f_{avg}^{-1}\right) r + \rho\tilde{\varepsilon}_i + z \tag{123}$$

The result on the covariance then follows trivially.

H.2 Result 4

Approximating first differences with derivatives, we obtain

$$\Delta Q_{i,t} - \Delta Z_t \approx -\sum_{j=0}^{\frac{T}{2}} \frac{2\pi j}{T} \left[\begin{array}{c} \sin(2\pi j t/T)\left(\rho\left[\left(f_{i,j}^{-1} - f_{avg,j}^{-1}\right) r_j + f_{i,j}^{-1}\tilde{\varepsilon}_{i,j}\right]\right) \\ +\cos(2\pi j t/T)\left(\rho\left[\left(f_{i,j}^{-1} - f_{avg,j}^{-1}\right) r_{j'} + f_{i,j}^{-1}\tilde{\varepsilon}_{i,j'}\right]\right) \end{array} \right] \tag{124}$$

where the approximation becomes a true equality as $T \to \infty$. Now if we furthermore use the approximations $f_{i,j_i^*}^{-1} - f_{avg,j_i^*}^{-1} \approx \bar{f}^{-1}/2$ and suppose that the exogenous supply process is small enough that it rarely causes a trader's demand to change signs, then we have

$$|\Delta Q_{i,t}| \approx |\Delta Z_t| + \omega_{j_i^*} \bar{f}^{-1} \rho \left| \begin{array}{c} \sin\left(\omega_{j_i^*} t\right)\left(r_{j_i^*} + \tilde{\varepsilon}_{i,j_i^*}\right) \\ +\cos\left(\omega_{j_i^*} t\right)\left(r_{j_i^{*\prime}} + \tilde{\varepsilon}_{i,j_i^{*\prime}}\right) \end{array} \right|. \tag{125}$$

H.3 Result 5

$$QV\{q_j\} \equiv \sum_{t=2}^{T} (Q_{i,t} - Q_{i,t-1})^2 \approx \sum_{t=2}^{T} \left(\sum_j \frac{2\pi j}{T} \left[\begin{array}{c} q_j \sin(2\pi jt/T) \\ +q_{j'} \cos(2\pi jt/T) \end{array} \right] \right)^2 \quad (126)$$

$$= \sum_{t=2}^{T} \sum_{j,k} \left(\frac{2\pi}{T} \right)^2 jk \left[\begin{array}{c} q_j \sin(2\pi jt/T)\, q_k \sin(2\pi kt/T) + q_{j'} \cos(2\pi jt/T)\, q_k \sin(2\pi kt/T) \\ q_j \sin(2\pi jt/T)\, q_{k'} \cos(2\pi kt/T) + q_{j'} \cos(2\pi jt/T)\, q_{k'} \cos(2\pi kt/T) \end{array} \right] \quad (127)$$

$$\approx \sum_{j,j'} (2\pi j)^2 T^{-1} q_j^2 \quad (128)$$

where the equality in the first line is approximate in assuming that $\cos(2\pi jt/T) - \cos(2\pi j(t-1)/T) \approx \frac{2\pi j}{T} \sin(2\pi jt/T)$ and the same for the differences in the sines. The third line uses the fact that sines of unequal frequencies are orthogonal (it is approximate because $t = 1$ is not included in the sum inserts the integral for $\sin^2$ and $\cos^2$, rather than the exact finite sums. All the approximations here are accurate for large T.

I Proofs of trading restriction results

I.1 Results 6 and 7

If trade by the investors is not allowed at certain frequencies, then obviously markets cannot clear at those frequencies when supply is inelastic. In this section we therefore first solve the model for the case with an upward sloping supply curve and then analyze the effect of eliminating trade on asset prices and returns.

I.1.1 Equilibrium with elastic supply

We now assume that there is exogenous supply on each date of

$$Z_t = \tilde{Z}_t + kP_t \quad (129)$$

where k is a constant determining the slope of the supply curve. One could imagine allowing k to differ across frequencies, which would be equivalent to allowing supply to depend on prices on multiple dates (intuitively, maybe supply increases by more when prices have been persistently high than when they are just temporarily high). Here, though, we simply leave k constant across frequencies. Multiplying by Λ' yields

$$z_j = \tilde{z}_j + kp_j \quad (130)$$

Solving the inference problem as before, we obtain

$$\tau_i \equiv Var\left[d \mid y_i, p\right]^{-1} \tag{131}$$

$$= \frac{a_1^2}{a_2^2} f_{\tilde{Z}}^{-1} + f_i^{-1} + f_D^{-1} \tag{132}$$

and

$$E\left[d \mid y_i, p\right] = \tau_i^{-1}\left(f_i^{-1} y_i + \frac{a_1}{a_2^2} f_{\tilde{Z}}^{-1} p\right) \tag{133}$$

I.1.2 Demand and equilibrium

The investors' demand curves are again

$$q_i = \rho_i\left(f_i^{-1} y_i + \left(\frac{a_1}{a_2^2} f_{\tilde{Z}}^{-1} - \tau_i\right) p\right)$$

Summing up all demands and inserting the guess for the price process yields

$$\tilde{z} + k\left(a_1 d - a_2 \tilde{z}\right) = \int_i \rho\left(f_i^{-1} d + \left(\frac{a_1}{a_2^2} f_{\tilde{Z}}^{-1} - \tau_i\right)\left(a_1 d - a_2 \tilde{z}\right)\right) di \tag{134}$$

Matching coefficients yields

$$\int_i \rho\left(\frac{a_1}{a_2^2} f_{\tilde{Z}}^{-1} - \tau_i\right) di = -a_2^{-1}\left(1 - k a_2\right) \tag{135}$$

$$\int_i \rho f_i^{-1} + \rho\left(\frac{a_1}{a_2^2} f_{\tilde{Z}}^{-1} - \tau_i\right) a_1 di = k a_1 \tag{136}$$

Combining those two equations, we have

$$\rho f_{avg}^{-1} = a_1\left(k + a_2^{-1}\left(1 - k a_2\right)\right) \tag{137}$$

$$= \frac{a_1}{a_2} \tag{138}$$

$$a_1 = \frac{f_{avg}^{-1} + \left(\rho f_{avg}^{-1}\right)^2 f_{\tilde{Z}}^{-1}}{\tau_{avg} + \rho^{-1} k} \tag{139}$$

$$= \frac{\tau_{avg} - f_D^{-1}}{\tau_{avg} + \rho^{-1} k} \tag{140}$$

$$a_2 = \frac{a_1}{\rho f_{avg}^{-1}} \tag{141}$$

I.1.3 Utility

As before, the contribution to optimized utility from frequency j is

$$\left((1 - a_{1,j})^2 f_{D,j} + \frac{a_{1,j}^2}{(\rho f_{avg}^{-1})^2} f_{Z,j} \right) \tau_{i,j} \tag{142}$$

Furthermore,

$$\begin{aligned}
(1 - a_{1,j})^2 f_{D,j} + \frac{a_{1,j}^2}{\left(\rho f_{avg,j}^{-1}\right)^2} f_{\tilde{Z},j} &= \left(\frac{\rho^{-1} k + f_D^{-1}}{\tau_{avg} + \rho^{-1} k} \right)^2 f_{D,j} + \rho^{-2} f_{avg,j}^2 \left(\frac{\tau_{avg} - f_D^{-1}}{\tau_{avg} + \rho^{-1} k} \right)^2 f_{\tilde{Z},j} \\
&= \left(\tau_{avg} + \rho^{-1} k\right)^{-2} \left(\left(\rho^{-1} k + f_D^{-1}\right)^2 f_{D,j} + \rho^{-2} f_{avg,j}^2 \left(\left(\rho f_{avg}^{-1}\right)^2 f_{\tilde{Z}}^{-1} + f_{avg}^{-1} \right)^2 f_{\tilde{Z},j} \right) \\
&= \left(\tau_{avg} + \rho^{-1} k\right)^{-2} \left(\left(\rho^{-1} k + f_D^{-1}\right)^2 f_{D,j} + \rho^{-2} \left(\rho^2 f_{avg}^{-1} f_{\tilde{Z}}^{-1} + 1 \right)^2 f_{\tilde{Z},j} \right)
\end{aligned}$$

So

$$E_{-1}\left[U_{i,0}\right] = \frac{1}{2} T^{-1} \sum_{j=0}^{T-1} \left[\begin{array}{c} \left(\tau_{avg,j} + \rho^{-1} k\right)^{-2} \left(\left(\rho^{-1} k + f_{D,j}^{-1}\right)^2 f_{D,j} + \rho^{-2} \left(\rho^2 f_{avg,j}^{-1} f_{\tilde{Z},j}^{-1} + 1 \right)^2 f_{\tilde{Z},j} \right) \\ \times \left(\left(\rho f_{avg,j}^{-1}\right)^2 f_{Z,j}^{-1} + f_{i,j}^{-1} + f_{D,j}^{-1} \right) \end{array} \right] - \frac{1}{2} \tag{143}$$

When there are no active investors and just exogenous supply, we have

$$0 = \tilde{z}_j + k p_j \tag{144}$$

$$p_j = -k^{-1} \tilde{z}_j \tag{145}$$

$$r_j = d_j - k^{-1} \tilde{z}_j \tag{146}$$

We then have

$$f_R = f_D + \frac{f_Z}{k^2} \tag{147}$$

$$f_{R,0} = f_{D,j} + \frac{f_{\tilde{Z},j}}{\left(k + \rho f_{D_j}^{-1}\right)^2} \tag{148}$$

I.2 Informativeness for moving averages of D_t

If a person is making decisions based on estimates of fundamentals from prices and they are worried that prices are contaminated by high-frequency noise, a natural response would be to examine an average of fundamentals and prices over time. For averages of fundamentals, we have the following:

Result 8 *The variance of an estimate of the average of fundamentals over dates t to $t + n - 1$*

conditional on observing the vector of prices, P, is

$$Var\left(\frac{1}{n}\sum_{m=0}^{n-1} D_{t+m} \mid P\right) = \frac{1}{nT}\sum_{j,j'} F_n(\omega_j)\,\bar{\tau}_j^{-1} \tag{149}$$

$$\text{where } F_n(\omega_j) \equiv \frac{1}{n}\frac{1-\cos(n\omega_j)}{1-\cos(\omega_j)} \tag{150}$$

and $\bar{\tau}_j = Var(d_j \mid p)$ as given by Remark 6.

F_n is the Fejér kernel. $F_1 = 1$, and as n rises, the mass of the Fejér kernel migrates towards the origin. That is, it places progressively less mass on high frequencies and more on low frequencies (it always integrates to 1). Specifically,

$$\frac{1}{T}\sum_{j,j'} F_n(\omega_j) = 1 \tag{151}$$

$$\lim_{n\to\infty} F_n(\omega) = 0 \text{ for all } \omega \neq 0 \tag{152}$$

The total weight allocated across the frequencies always sums to 1, and as n rises, the mass becomes allocated eventually purely to frequencies local to zero.

This result shows that the informativeness of prices for moving averages of fundamentals places relatively more weight on low- than high-frequency informativeness. So even if prices have little or no information at high frequencies – $\bar{\tau}_j$ is small for large j, there need not be any degradation of information about averages of fundamentals over multiple periods, as they depend primarily on precision at lower frequencies (smaller values of j).

The top panel of figure 4 plots the Fejér kernel, F_n, for a range of values of n. One can see that even with $n = 2$, the weight allocated to frequencies above the cutoff of $\omega = 3$ that we use in the example in figure 3 is close to zero. As n rises higher, the weight falls towards zero at a progressively wider range of frequencies. Equation (149) therefore shows that while a reduction in precision at high frequencies due to trading restrictions will reduce the informativeness of prices about fundamentals on any single date, it has quantitatively small effects on the informativeness of prices for fundamentals over longer periods.

Moving averages of fundamentals depend less on the precise high-frequency details of the world, so when high-frequency information is reduced, we would not expect to see a reduction in the informativeness of prices for moving averages. More concretely, going back to our example of oil futures, when high-frequency trade is not allowed, prices become noisier, making it more difficult to obtain an accurate forecast of the spot price of oil at some specific moment in the future. If one is interested in the average of spot oil prices over a year, on the other hand, then we would expect futures prices to remain informative, even when high-frequency trade is restricted.

I.2.1 Proof of result 8

We have

$$D \mid Y, P \sim N\left(\bar{D}, \Lambda diag\left(\tau_0^{-1}\right)\Lambda'\right) \quad (153)$$

where τ_0 is a vector of frequency-specific precisions conditional on prices. Now consider some average over D, $F'D$, where F is a column vector. Then

$$\begin{aligned} Var\left(D_t\right) &= 1_t'\Lambda diag\left(\tau_0^{-1}\right)\Lambda' 1_t & (154)\\ &= \left(\Lambda' 1_t\right)' diag\left(\tau_0^{-1}\right)\left(\Lambda' 1_t\right) & (155)\\ &= \sum_{j,j'} \lambda_{t,j}^2 \tau_{0,j}^{-1} & (156)\\ &= \lambda_{t,0}^2\tau_{0,0}^{-1} + \lambda_{t,\frac{T}{2}}^2\tau_{0,0}^{-1} + \sum_{n=1}^{\frac{T}{2}-1}\left(\lambda_{t,n}^2 + \lambda_{t,n'}^2\right)\tau_{0,n}^{-1} & (157)\end{aligned}$$

where 1_t is a vector equal to 1 in its tth element and zero elsewhere and $\lambda_{t,j}$ is the jth trigonometric transform evaluated at t, with

$$\begin{aligned} \lambda_{t,j} &= \sqrt{2/T}\cos\left(2\pi j\left(t-1\right)/T\right) & (158)\\ \lambda_{t,j'} &= \sqrt{2/T}\sin\left(2\pi j\left(t-1\right)/T\right) & (159)\\ \lambda_{t,0} &= \sqrt{1/T} & (160)\\ \lambda_{t,\frac{T}{2}} &= \sqrt{1/T}\cos\left(\pi\left(t-1\right)\right) = \sqrt{1/T}\left(-1\right)^{t-1} & (161)\end{aligned}$$

More generally, then

$$\begin{aligned} Var\left(\frac{1}{s}\sum_{m=0}^{s-1} D_{t+m}\right) &= \frac{1}{s^2}\left(\sum_{m=0}^{s-1} 1_{t+m}\right)' \Lambda diag\left(\tau_0^{-1}\right)\Lambda'\left(\sum_{m=0}^{s-1} 1_{t+m}\right) & (162)\\ &= \frac{1}{s^2}\left(\sum_{m=0}^{s-1}\lambda_{t+m,0}\right)^2\tau_{0,0}^{-1} + \frac{1}{s^2}\left(\sum_{m=0}^{s-1}\lambda_{t+m,\frac{T}{2}}\right)^2\tau_{0,\frac{T}{2}}^{-1} & (163)\\ &\quad +\frac{1}{s^2}\sum_{n=1}^{\frac{T}{2}-1}\left[\left(\sum_{m=0}^{s-1}\lambda_{t+m,n}\right)^2 + \left(\sum_{m=0}^{s-1}\lambda_{t+m,n'}\right)^2\right]\tau_{0,n}^{-1} & (164)\end{aligned}$$

where $\tau_{0,n}$ is the frequency-n element of τ_0. For $0 < n < \frac{T}{2}$

$$\left(\sum_{m=0}^{s-1}\lambda_{t+m,n}\right)^2 + \left(\sum_{m=0}^{s-1}\lambda_{t+m,n}\right)^2 = \sum_{m=0}^{s-1}\sum_{k=0}^{s-1}\frac{2}{T}\left[\begin{array}{c}\cos\left(2\pi n\left(t+m-1\right)/T\right)\cos\left(2\pi n\left(t+k-1\right)/T\right)\\ +\sin\left(2\pi n\left(t+m-1\right)/T\right)\sin\left(2\pi n\left(t+k-1\right)/T\right)\end{array}\right] \quad (165)$$

Now note that

$$2\cos\left(x\right)\cos\left(y\right) + 2\sin\left(x\right)\sin\left(y\right) = 2\cos\left(x-y\right) \quad (166)$$

So we have

$$\left(\sum_{m=0}^{s-1}\lambda_{t+m,n}\right)^2+\left(\sum_{m=0}^{s-1}\lambda_{t+m,n}\right)^2 = \frac{2}{T}\sum_{m=0}^{s-1}\sum_{k=0}^{s-1}\cos\left(\frac{2\pi n}{T}(m-k)\right) \tag{167}$$

$$= 2\frac{s}{T}\sum_{m=-(s-1)}^{s-1}\frac{s-|m|}{s}\cos\left(\frac{2\pi n}{T}m\right) \tag{168}$$

$$= 2\frac{s}{T}F_s\left(\frac{2\pi n}{T}\right) \tag{169}$$

$$= \frac{2}{T}\frac{1-\cos\left(s\frac{2\pi n}{T}\right)}{1-\cos\left(\frac{2\pi n}{T}\right)} \tag{170}$$

where F_s denotes the sth-order Fejér kernel. Note that when $s=T$, the above immediately reduces to zero, since $\cos(2\pi n)=0$. That is the desired result, as an average over all dates should be unaffected by fluctuations at any frequency except zero.

For $n=0$,

$$\left(\sum_{m=0}^{s-1}f_{t+m,0}\right)^2 = \left(\sum_{m=0}^{s-1}\sqrt{1/T}\right)^2 \tag{171}$$

$$= \left(s\frac{1}{T^{1/2}}\right)^2 \tag{172}$$

$$= \frac{s}{T}F_s(0) \tag{173}$$

Since $F_s(0)=s$ (technically, this holds as a limit: $\lim_{x\to 0}F_s(x)=s$).

For $n=\frac{T}{2}$,

$$\left(\sum_{m=0}^{s-1}f_{t+m,\frac{T}{2}}\right)^2 = \frac{1}{T}\left(\sum_{m=1}^{s}(-1)^m\right)^2 = \begin{cases}\frac{1}{T}\text{ for odd }s\\ 0\text{ otherwise}\end{cases} \tag{174}$$

$$= \frac{s}{T}\frac{1}{s}\left(\frac{\sin(s\pi/2)}{\sin(\pi/2)}\right)^2 = \frac{s}{T}F_s(\pi) \tag{175}$$

So we finally have that

$$Var\left(\frac{1}{s}\sum_{m=0}^{s-1}D_{t+m}\right) = \frac{1}{sT}\sum_{j,j'}F_s(\omega_j)\,\tau_{0,j}^{-1} \tag{176}$$

J Costly learning about prices

J.1 Generic result: no learning from prices

Lemma 8 *Assume that learning from prices is costly. At that at time -1, if agent i decides to infer information from prices, then their capacity constraint is:*

$$Tr(f_i^{-1} + f_P^{-1}) \leq \bar{f}^{-1},$$

where f_P^{-1} is inverse of the variance-covariance matrix of signals contained in prices, and f_i^{-1} is the variance-covariance of the private signals of agent i. On the other hand, if agent i decides not to infer information from prices, then his capacity constraint is:

$$Tr(f_i^{-1}) \leq \bar{f}^{-1}.$$

Then, agents always prefer not to learn from prices.

Proof. If agent i has decided not to learn from prices, then at time 0, their posterior distribution over d is:

$$\begin{aligned} d|y_i &\sim N\left(\mu(y_i), \tau_i^{-1}\right) \\ \tau_i^{NP} &= f_D^{-1} + f_i^{-1} \\ \mu(y_i) &= (\tau_i^{NP})^{-1} f_i^{-1} y_i \end{aligned} \tag{177}$$

Agent i still observes prices; their first-order condition leads to the demand schedule:

$$q_i = \rho \tau_i^{NP} \left(\mu(y_i) - p\right).$$

His time-0 utility is:

$$U_{0,i}^{NP}(y_i; p) \quad = \tfrac{1}{2T} \left(\mu(y_i) - p\right)' \tau_i^{NP} \left(\mu(y_i) - p\right). \tag{178}$$

Since τ_i^{NP} is symmetric, this implies:

$$E_{-1,i}\left[U_{0,i}^{NP}\right] \quad = \tfrac{1}{2T} tr(\tau_i^{NP} V_i^{NP}) + \tfrac{1}{2T} (\mu_i^{NP})' \tau_i^{NP} \mu_i^{NP}, \tag{179}$$

where as before:

$$\begin{aligned} \mu_i^{NP} &= E_{-1,i}\left[\mu(y_i) - p\right] \\ V_i^{NP} &= Var_{-1,i}\left[\mu(y_i) - p\right] \end{aligned} \tag{180}$$

As before, because all fundamentals are mean 0, $\mu_i = 0$. Moreover, by the law of total variance:

$$V_i = \underbrace{Var_{-1}\left[d - p\right]}_{\equiv V_{-1}} - (\tau_i^{NP})^{-1}$$

Therefore,

$$
\begin{aligned}
E_{-1,i}\left[U_{0,i}^{NP}\right] &= \tfrac{1}{2T}tr(\tau_i^{NP}V_i) \\
&= \tfrac{1}{2T}tr(\tau_i^{NP}V_{-1}) - \tfrac{1}{2T}tr(I) \\
&= \tfrac{1}{2T}tr\left(f_D^{-1}V_{-1}\right) - \tfrac{1}{2T}tr(I) + \tfrac{1}{2T}tr(f_i^{-1}V_{-1})
\end{aligned} \tag{181}
$$

The time-(-1) attention allocation problem of such an agent is therefore:

$$
\begin{aligned}
U_{-1,i}^{NP}\left(f_{avg}^{-1}\right) \quad & = -\tfrac{1}{2} + \tfrac{1}{2T}tr\left(f_D^{-1}V_{-1}\right) + \quad \frac{1}{2T}\max_{f_i^{-1}} \quad tr(f_i^{-1}V_{-1}) \\
\text{s.t.} \quad & f_{i,j}^{-1} \geq 0 \quad \forall j \in [0, ..., T-1] \\
& tr(f_i^{-1}) \leq \bar{f}^{-1}
\end{aligned} \tag{182}
$$

For an agent who does learn from prices (but shares the other agent's ex-ante distribution over p and d, summarized by V_{-1}), the attention allocation problem has already been derived; it is given by:

$$
\begin{aligned}
U_{-1,i}\left(f_{avg}^{-1}\right) \quad & = -\tfrac{1}{2} + \tfrac{1}{2T}tr\left(\left(f_D^{-1} + f_P^{-1}\right)V_{-1}\right) + \quad \frac{1}{2T}\max_{f_i^{-1}} \quad tr(f_i^{-1}V_{-1}) \\
\text{s.t.} \quad & f_{i,j}^{-1} \geq 0 \quad \forall j \in [0, ..., T-1] \\
& tr(f_i^{-1} + f_P^{-1}) \leq \bar{f}^{-1}
\end{aligned} \tag{183}
$$

Since f_i^{-1} is diagonal, $f_i^{-1} \to tr(f_i^{-1}V_{-1})$ can be thought of as a linear map on R^T. By the Riesz representation theorem, there is $\lambda \in R^T$ such that $\forall f_i^{-1}$, $tr(f_i^{-1}V_{-1}) = \sum_{j=0}^{T-1} f_{i,j}^{-1}\lambda_j$. Let $\tilde{\lambda}$ denote the element-wise maximum of λ. Note, in particular, that:

$$
tr(f_P^{-1}V_{-1}) = \sum_{j=0}^{T-1} f_{P,j}^{-1}\lambda_j.
$$

Moreover, after optimization, not learning through prices yields utility:

$$
U_{-1,i}^{NP}\left(f_{avg}^{-1}\right) = -\frac{1}{2} + \frac{1}{2T}tr\left(f_D^{-1}V_{-1}\right) + \frac{1}{2T}\tilde{\lambda}\bar{f}^{-1}.
$$

Learning through prices yields utility:

$$
U_{-1,i}\left(f_{avg}^{-1}\right) = -\frac{1}{2} + \frac{1}{2T}tr\left(\left(f_D^{-1} + f_P^{-1}\right)V_{-1}\right) + \frac{1}{2T}\tilde{\lambda}\left(\bar{f}^{-1} - tr(f_P^{-1})\right)
$$

The difference between the two is:

$$
\begin{aligned}
U_{-1,i}^{NP}\left(f_{avg}^{-1}\right) - U_{-1,i}\left(f_{avg}^{-1}\right) \quad & = \tfrac{1}{2T}\tilde{\lambda}tr(f_P^{-1}) - \tfrac{1}{2T}tr\left(f_P^{-1}V_{-1}\right) \\
& = \tfrac{1}{2T}\tilde{\lambda}tr\left(f_P^{-1}\right) - \tfrac{1}{2T}\textstyle\sum_{j=0}^{T-1} f_{P,j}^{-1}\lambda_j \\
& \geq 0
\end{aligned} \tag{184}
$$

Therefore, the agent always prefer not to learn from prices. ■

J.2 The equilibrium when agents do not learn about prices

Guess:

$$p = a_3 d - a_4 z$$

with a_3, a_4 diagonal matrices of size $T \times T$. Straightforward derivations lead to:

$$\begin{aligned}
a_3 &= I - (\tau_{avg} + kI)^{-1} \left(f_D^{-1} + kI\right) \\
&= (\tau_{avg} + kI)^{-1} f_{avg}^{-1} \\
a_4 &= \tfrac{1}{\rho} a_3 f_{avg} \\
&= \tfrac{1}{\rho} (\tau_{avg} + kI)^{-1} \\
\tau_{avg} &= f_{avg}^{-1} + f_D^{-1} \\
\tau_i &= f_i^{-1} + f_D^{-1}
\end{aligned} \tag{185}$$

Moreover, expected utility is given by:

$$\begin{aligned}
E_{-1,i}\left[U_{0,i}^{NP}\right] &= \mathcal{C}^{NP} + \tfrac{1}{2T} tr(V_{-1}^{NP} f_i^{-1}) \\
V_{-1}^{NP} &= f_D \left((I + kf_D)^2 + \tfrac{f_Z f_D}{\rho^2}\right) (I + kf_D + f_D f_{avg}^{-1})^{-2} \\
\mathcal{C}^{NP} &= \tfrac{1}{2T} tr\left(f_D^{-1} V_{-1}^{NP}\right) - \tfrac{1}{2}
\end{aligned} \tag{186}$$

K Calibration

$\bar{f}^{-1} = 0.01$

$T = 1000$

$f_D(\omega) = \frac{1}{4}\left|1 - \frac{1}{2}e^{i\omega}\right|^{-2} + 1 - .55\cos(2\omega) + \frac{7}{16}\left|1 + \frac{1}{2}e^{i\omega}\right|^{-2}$

$f_Z(\omega) = \frac{1}{10}\left|1 - \frac{1}{2}e^{i\omega}\right|^{-2}$

$\rho = 1$

Figure 1: Optimal information acquisition and waterfilling

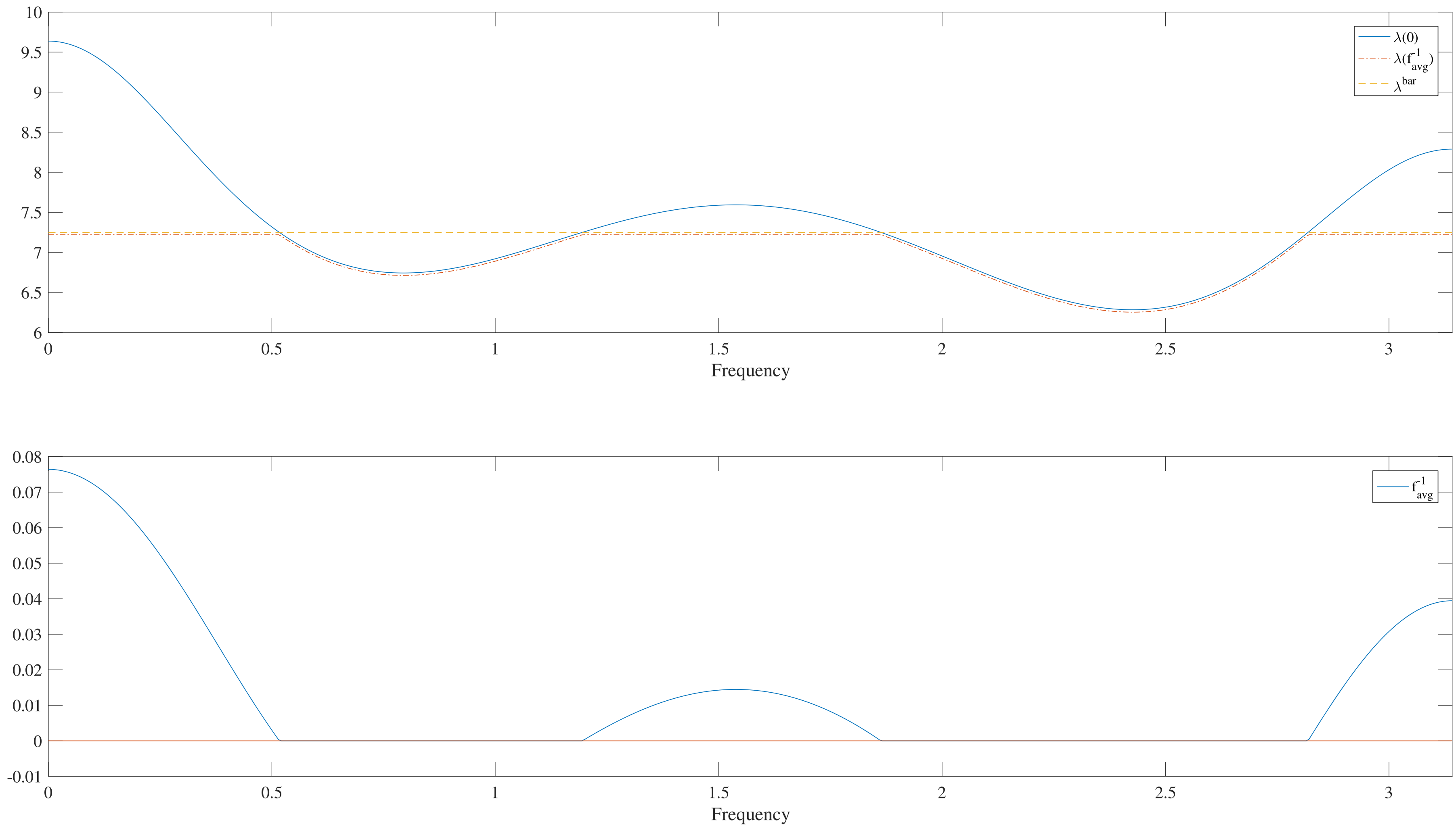

Figure 2: Example investor's demand

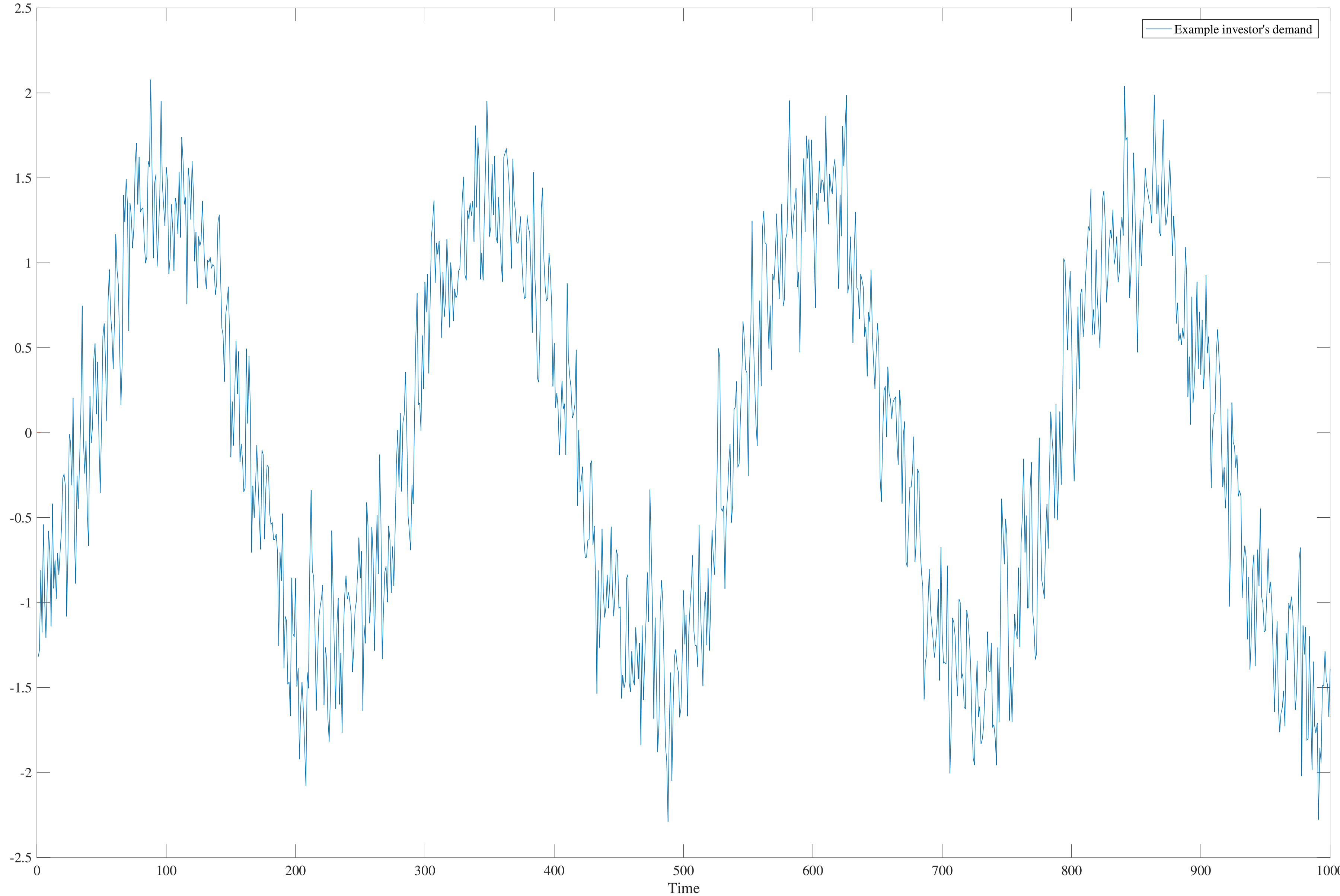

Figure 3: Effects of restricting high-frequency trade

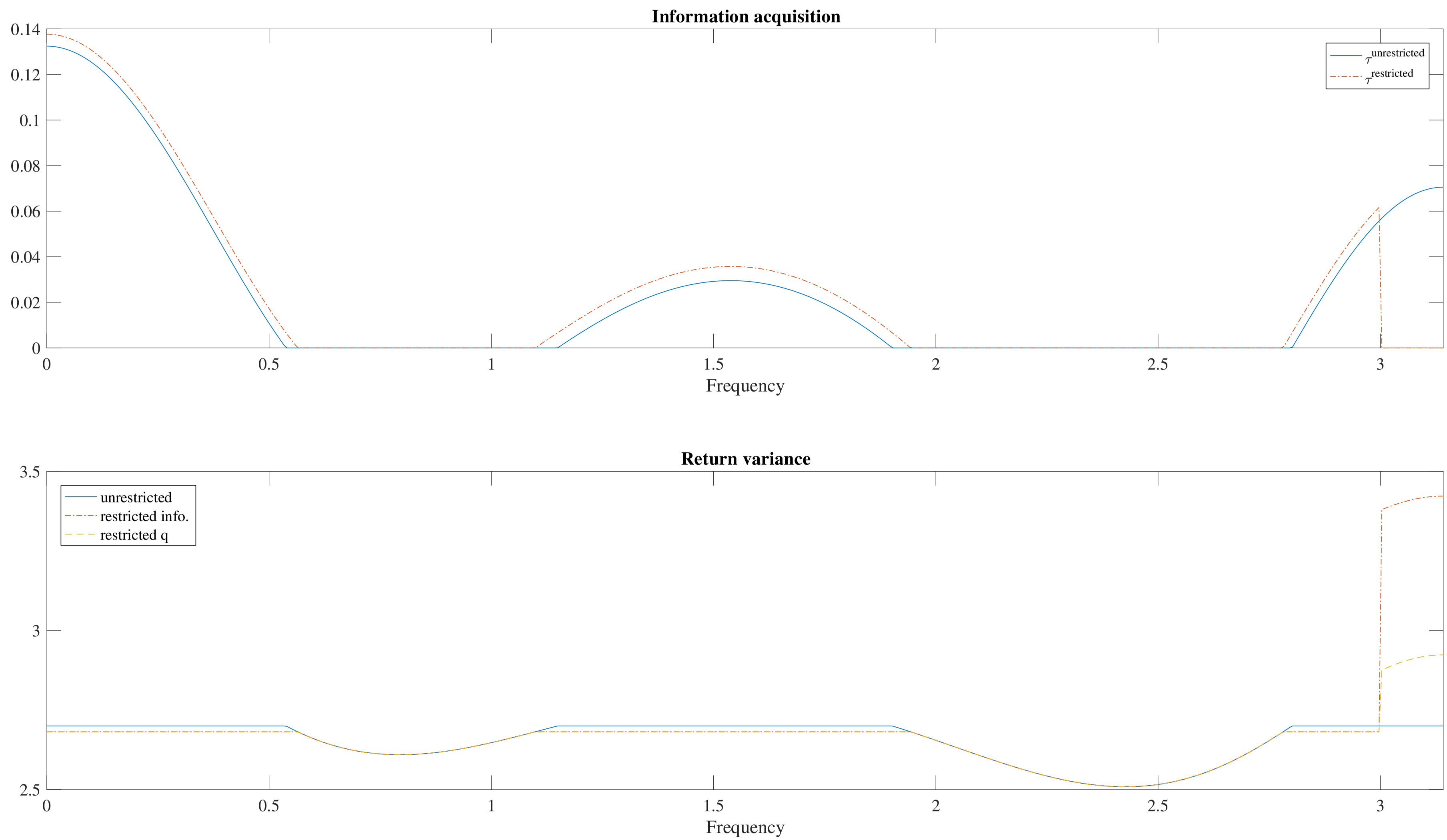

Figure 4: Weights across frequencies and time

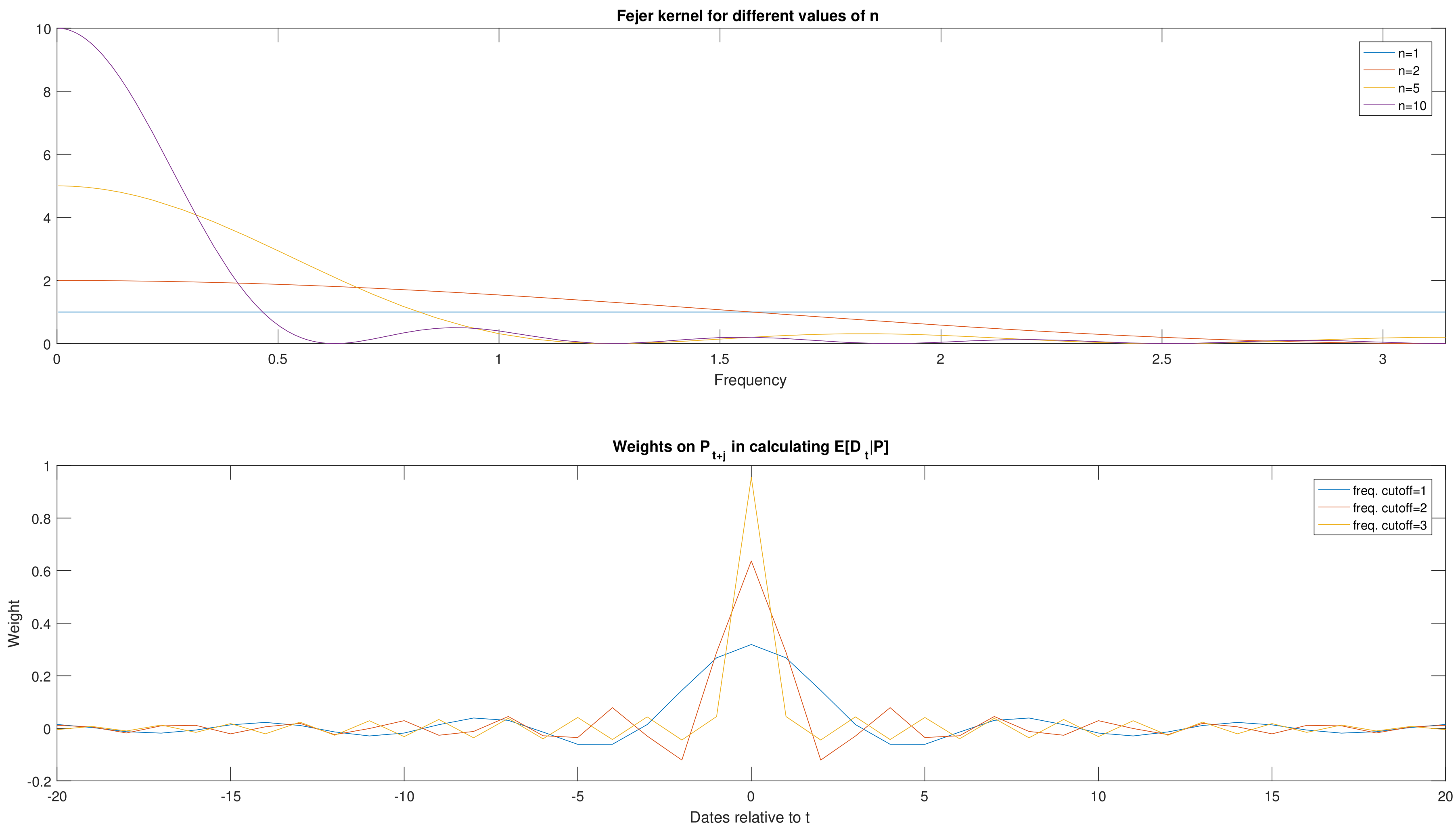

Figure 5: Persistence of the churn rate over time

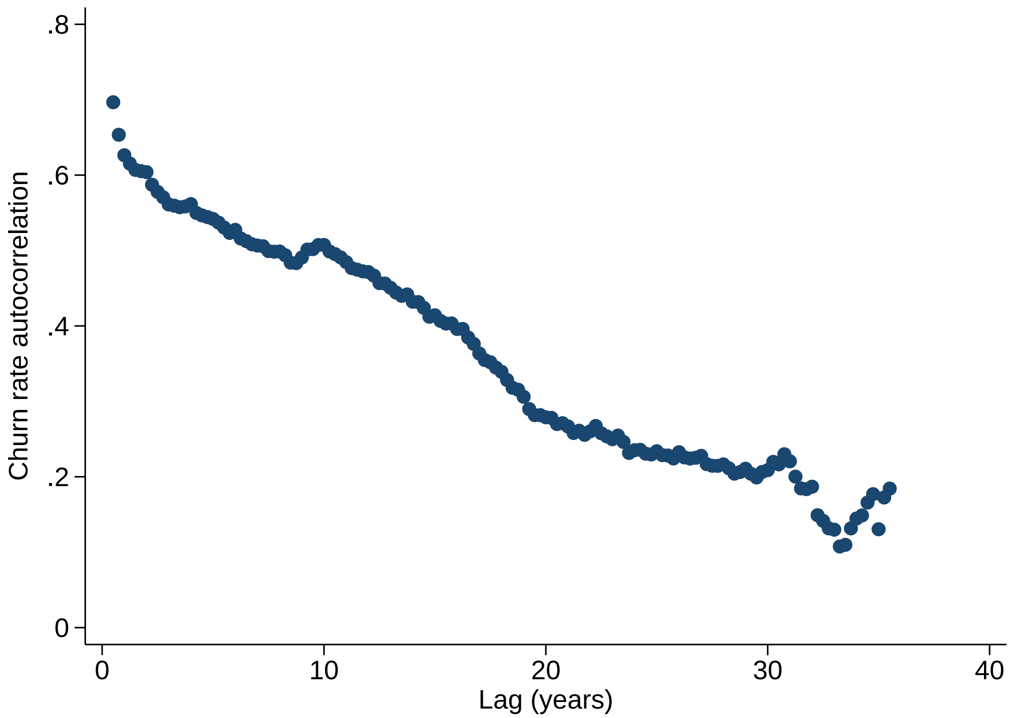

Figure 6: Out-performance of institution holdings at different horizons